EXPERIMENTS IN IDENTITY ECONOMICS

A Dissertation

by

NOAH CRAIG BACINE

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Catherine C. Eckel
Committee Members,	Alexander L. Brown
	Marco Castillo
	Marco Palma
Head of Department,	Timothy Gronberg

May 2019

Major Subject: Economics

ABSTRACT

This dissertation furthers our understanding of the influence of identity on behavior using experimental methods. My first project investigates the influence of social norms on behavior by testing for the presence of a gender gap in competitive preferences across two populations that vary in culture. Using a laboratory experiment in which subjects select among a competitive and non-competitive payment scheme, we find a stark contrast across the two populations with one exhibiting the previously found gender gap in competitive preferences while the other exhibits no such gap. Our results suggest that nurture in the form of social norms may explain the persistence of the gender disparity in labor market outcomes rather than nature.

My second project looks at group behavior more generally by studying the mechanism underlying in-group bias in trust. Using a laboratory experiment that separately identifies the ways in which the decision to trust differs from a similarly risky decision, we study how individuals' willingness to trust depends on the identity of their counterpart. We find that individuals exhibit little to no preference over relative earnings whether the subject's counterpart is a member of their group or another group. On the other hand, individuals are significantly less likely to take a risk when its outcome is determined by a member of another group as opposed to a member of their own group. Our results imply that the in-group bias in trust is the result of differences in the perceived emotional cost of betrayal rather than altruistic preferences towards the in-group.

My final project returns to studying gender norms but does so in the context of social interaction. We study the strength of gender norms in low promotability task allocation and completion through the introduction of heterogeneous costs into a laboratory setting. We find that explicit cost differences produce an overall gender gap that is similar in magnitude to previous

studies, but that the gap among low cost individuals is smaller. In an additional treatment, we find that when managers who request another participant to complete the task are introduced that the inclusion of explicit cost differences crowds out the previously found gender bias. Taken together, our results suggest that gender beliefs play a large role in the gender disparity in low promotability task allocation but that making managers aware of opportunity cost differences among their employees may alleviate this trend.

My three projects illustrate that identity is influential in the development of individual preferences. Individuals recognize and conform to the behavioral prescriptions of the groups to which they belong leading to large differences in behavior between and across groups. My results suggest that a deeper understanding of the influence of norms and identity-based behavioral prescriptions is necessary to develop a deeper understanding of heterogeneity in behavior.

DEDICATION

To my mother, Ilene Bacine and my sister, Jessica Bacine:

For the unconditional love and support always.

To my family and friends:

For never letting me forget what matters.

To my father, Russell Bacine:

For being my hero and my role model while teaching me to be even more. I hope that my work can contribute to creating the world that you taught me to believe in.

ACKNOWLEDGEMENTS

This dissertation is the culmination of my development as a scholar and as a mentor over the past six years. Although it is product of my efforts, it is also a reflection of the strength given to me by mentors, peers, friends, and family. The amount of space it would take to properly thank each person who has contributed to my success would be much longer than the dissertation itself, but I will do my best to provide a brief glimpse into how thankful I am that each and every one of you has been with me every step of the way.

By far, the greatest support I have received is from my advisor, Catherine C. Eckel, whose patience and guidance shaped me into the researcher I am today. Your mentorship opened my eyes to the world of behavioral economics, and guided me along the path to developing my own research. You have sacrificed your resources to safeguard my success, to ensure my ability to run experiments, and to provide opportunities to present my research all over the world. You introduced me to a number of individuals who have in turn welcomed and encouraged me as a young academic. You will always be my definition of the word mentor.

The other members of my committee have been an incredible source of support and advice as well. I appreciate the time that each of you took to provide guidance and advice about developing my own research agenda. I give a special thanks to Marco Palma. You are the most supportive academic I have ever met.

There are many individuals outside of my university who have been integral to my development as well. I thank Dr. Gbenga Ojumu for being a team player and helping me as a fellow instructor in uncharted territory. I thank Dr. Rick Wilson for allowing me to use his laboratory at Rice University for my first experiment. I thank the various students, especially those

who took a directed studies course with me, for providing the fuel to push my own research forward and the appreciation to continue to act as a mentor.

I could never have finished this journey without my friends and family. My peers provided a community that supported me in my tribulations and celebrated with me in my triumphs. My friends were always ready with a sympathetic ear and propped me up when I lost my balance. My family provided the unconditional love and support that made me believe that this feat could be accomplished. The gift that you all have given me is ineffable and all I can provide in return is my eternal thanks.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of my advisor, Dr. Catherine C. Eckel of the Department of Economics, Dr. Alexander L. Brown of the Department of Economics, Dr. Marco Castillo of the Department of Economics, and Dr. Marco Palma of the Department of Agricultural Economics.

The first essay (Section 2) in my dissertation grew out of a directed studies course that was co-taught along with my advisor, Dr. Catherine C. Eckel, and Dr. Gbenga Ojumu of Prairie View A&M University in the spring of 2017. The student participants in the course were Romaine Knight, Nathaniel Mezgebe, Shea Niswanger, Bethany Patterson, Rebekah Preston, Antoinette Williams, and Matthew VanDivier. During the course, the students participated in designing and conducting an experimental study in conjunction with the instructors. The students were responsible for initial data collection and analysis as part of their requirements for the course. My chapter incorporates this data, and extends the original study we developed during the course. In the year following the course, I ran additional sessions to increase our sample size and devoted special attention to acquiring a sample of black students at Texas A&M University that serve as an important control in understanding our results. I also conducted additional data analysis, bringing to bear more sophisticated statistical tools that are in line with previously published work.

The data used in the second essay (Section 3) was collected with the permission and help of Dr. Rick Wilson of Rice University. Dr. Wilson helped me get access to the Rice University subject pool and made his lab available for conducting the sessions.

The third essay (Section 4) also grew out of a directed studies course. This course was co-taught along with my advisor, Dr. Catherine C. Eckel, in the fall of 2016. Savannah Hedge, Kaitlyn Kellermeyer, Madeline Mitchell, Ryan Mulvihill, and Adam Williams were the student participants. Under my leadership, we developed the preliminary research idea together, and the students were involved in the initial data collection and analysis as part of their requirements for the directed studies course. I have further developed and extended this research by designing and conducting an additional treatment that was not part of the course project. Additionally, I redid and extended the analysis to account for potential collinearity among observations from the study.

All other work for the dissertation was completed by the student, in collaboration with Dr. Catherine C. Eckel of the Department of Economics.

Funding Sources

Graduate study was supported by a fellowship from Texas A&M University.

The first essay was supported by a grant from the National Science Foundation (SES-1530796). Additional research funding was provided by the Department of Economics, the Private Enterprise Research Center at Texas A&M University, and by a Ray C. Battalio Research Fellowship.

NOMENCLATURE

DOSPERT	Domain-Specific Risk-Taking
ERL	Economic Research Laboratory
ISIS	Islamic State in Iraq and Syria
MAP	Minimal Acceptable Probability
NLSY	National Longitudinal Survey of Youth
OLS	Ordinary Least Squares
ORSEE	Online Recruitment System for Economic Experiments
PVRL	Prarie View Research Laboratory
VD	Volunteer's Dilemma
ZTREE	Zurich Toolbox for Readymade Economic Experiments

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. GENERAL INTRODUCTION

This dissertation is comprised of three projects which study the influence of identity on behavior. Within this work, an identity is defined as the recognition of membership to a social group by an individual as well as others. Although standard utility theory would suggest that rational individuals would ignore such considerations, Akerloff and Kranton's (2000) theoretical work and research across disciplines has indicated this is not the case. The work that comprises this dissertation contributes to our understanding of identity through investigating identity-based social interaction and the importance of group-specific norms in the development of preferences and their resulting behavior.

The first project (section 2) studies the role of norms in the formation of economic disparity across groups. In particular, we look at how differences in gender norms may explain the persistence of the gender gap in labor market outcomes. Niederle and Vesterlund (2007) find a significant gender gap in competitive preferences and argue that this difference may explain the significant gender gap in choice of occupation. However, it is unclear what causes the observed gender gap in competitive preferences. We join others in seeking to understand what may explain this difference in preferences.

To do so, we study two populations that differ in culture to examine the effect of social norms on the expression of competitive preferences. We replicate the procedures of Niederle and Vesterlund (2007) in which subjects complete a mathematics task under a competitive and non-competitive payment scheme before they are asked to select which payment scheme they would like to apply to an additional performance. We repeat this procedure with Black and Hispanic subjects who come from populations that differ in their gender norms to measure the effect of culture on the expression of competitive preferences.

We find a gap that is almost identical to the original study in the Hispanic sample but no gender gap in competitive preferences in the black population. After ruling out a variety of alternative explanations, we are left to conclude that these populations differ in a way that significantly effects the expression of competitive preferences. This suggests that nurture plays an important role in the previously observed gender gap in competitive preferences and that policies may effectively reduce the gap. Beyond this, our work suggests the importance and value of considering the multiple dimensions of identity in understanding preferences in behavior. In particular, we find that one identity can moderate the effect of another identity on behavior. Although this work is focused on the influence of identity on individual behavior, the second project (section 3) studies the effect of general group membership on social behavior.

Trust is an essential building block of economic activity and previous work has shown that individuals are biased towards interacting with members of their own group. Although there is an evolutionary argument for the existence of in-group bias, its existence in modernity often leads to inefficient behavior and in extreme cases, costly conflict. Previous research shows that in-group bias in trust exists (Falk and Zehnder, 2013), but it is unclear what drives this tendency. We contribute to this work by conducting a laboratory experiment using real randomly assigned identity to study the mechanism underlying in-group bias in trust.

To do so, we adapt the design first introduced by Bohnet and Zeckhauser (2004) which compares the decision to trust to a similarly risky decision. In the original experiment, subjects experience one of three treatments which vary the presence of a second individual who earnings depend on the decision of another and whether the outcome of a taking a risk is determined by another individual or nature. We adapt this design by introducing the within-subjects factor of group identity to investigate how the identity of the other party effects the decision to trust. Thus,

subjects express their desire to take a gamble when it also effects a member of their in-group or their out-group and the desire to trust when the outcome is determined by an in-group or an out-group member.

We find that subjects exhibit no significant preferences over relative earnings whether their partner is a member of their group or another group. On the other hand, we find that in-group bias in trust is the result of differences in the perceived emotional cost of betrayal (not having trust reciprocated by another individual). Subjects are significantly less willing to trust when the decision to trust entails a potential lack of reciprocity from a member of their out-group. We hypothesize that this is either because the desire to conform to group ideals by trusting the in-group outweighs the emotional cost of betrayal or that what constitutes a betrayal differs for in-group and out-group individuals. Our results shed light on the mechanism underlying in-group bias in trust providing greater insight into how to encourage positive intergroup interaction. In the final project (section 4), I revisit social interaction, but instead of studying general group norms, I return to studying the effects of group specific norms in a social setting.

Furthering our understanding of the persistence of the gender disparity in labor market outcomes, Babcock et al. (2017) investigate the existence of a gender gap in low promotability task completion. Using empirical and laboratory data, the authors find that women are more likely to be asked and to complete tasks with low promotability. We expand this work by investigating the importance of gender norms in this context by introducing explicit cost differences in volunteering.

To test the strength of gender norms in low promotability task completion, we repeat the experimental procedures of Babcock et al. (2017) with the addition of heterogeneous costs of volunteering. In the experiment, subjects are placed into groups of three and must choose whether

or not to pay a cost so the entire group is better off. Similar to the real-world incentives present in a low promotability task, subjects desire the benefit but prefer not to be the one to pay the cost. We find that the introduction of explicit cost differences into this context is more salient for men. The result is that the overall gender gap is similar but among low cost individuals is smaller than the original study while the opposite is true of high cost individuals. However, when managers are introduced, the gender gap is eliminated by managers making requests equally across genders. Although costs were ineffective in reducing the gap in volunteering, it did eliminate the gender gap in requests received to volunteer. Our results suggest that providing mangers with information regarding the opportunity costs of their employees may alleviate the empirically observed gender gap in low promotability task completion.

Together, my results suggest that social norms can have subtle yet powerful effects on individual behavior that can create large disparities in economic outcomes. Beyond individual behavior, my results illustrate that identity impacts how individuals interact with others leading to loss of economic inefficiency. In sum, my work illustrates that identity is a influential force that drives individuals' preferences and their resulting behavior.

2. GENDER, CULTURE, AND COMPETITION

2.1 Introduction

Although women have gained equality in some important areas, the gender gap in wages and upward mobility within firms have lagged. Women earn 80% of what a male in the same position earns on average and only 19 of the Fortune 500 companies have female Chief Executive Officers (Blau and Kahn, 2017). Historical explanations for the observed gender gap cite differences in ability and discrimination (Arrow, 1973; Polachek, 1981; Goldin and Rouse, 2000). Although these explanations may account for a small portion of the remaining gap, it is unlikely that either can fully account for its persistence.

In a recent review, Blau and Kahn (2017) claim that 60 percent of the gender gap in wages is due to differences in occupational choice. Occupational choice, in turn, is strongly correlated with competitive preferences (Buser et al. 2014; Buser et al. 2017). This suggests that an alternative explanation for the observed gap is gender differences in competitive preferences. Women only represent 30% of degree holders in science, technology, engineering and mathematics (STEM) fields, generally viewed as competitive majors, even though women now account for more than half of all college degree holders (Noonan, 2017). If firms award top positions competitively, then women's lower entry into competitive environments may explain the lack of female representation among top earning positions.

Testing this hypothesis in the lab, Niederle and Vesterlund (2007) find that men opt for a competitive pay scheme for performing a simple task twice as often as their female counterparts even though there are no differences in performance. Although a number of researchers have replicated this result in a number of settings using a variety of tasks (Niederle, 2015), the question of what causes the observed gap in competitive preferences is open.

Baron-Cohen and Benenson (2003) argue that the disparity in representation among the sciences results from the fact that men are innately better competitors. In favor of this argument, some researchers have found biological traits predict competitive preferences (Wozniak et al. 2014; Buser 2012; Hoffman and Gneezy, 2010). However, researchers have identified environments within which there exists no gap in competitive preferences (Booth and Nolen, 2010; Almås et al, 2015; Gneezy et al, 2009; Cassar et al. 2016). This suggests that biology alone cannot explain the gender gap observed in other populations. To understand nurture's role, researchers must investigate natural variation in women's experiences.[1]

Apart from biology, cultural norms may also play a significant role in shaping gender differences in competitiveness. Most studies exploring cultural correlates of competitiveness and risk-taking compare behavior of subjects across countries to show that the gender gap is not always present (Gneezy et al., 2009; Anderson et al. 2013; Cárdenas et al. 2012). Instead, we take advantage of cultural variation within the US by examining the competitiveness of black and Hispanic university students. We replicate the procedures of Niederle and Vesterlund (2007) and find robust evidence that culture plays a role in the formation of competitive preferences. Hispanic students exhibit a gap in preferences almost identical to the original paper while no such gap occurs in our black sample. Differences in ability cannot explain our results, as we detect no gender difference within either sample. Confidence and risk aversion explain a portion of the gap but neither can fully explain our results. Instead, using rich survey data, we find that differences in culture among the two samples predict our pattern of results. Our results add to the debate on whether nature or nurture explains the gender gap in competitive preferences by providing

[1] Ball et al. 2010 show that, physical prowess is correlated with risk-taking, a related trait to competitiveness. Their work also shows that others' perceptions of risk tolerance exaggerate the underlying differences related to the appearance of strength. This suggests that stereotypes about risk taking or competitiveness may lead people to expect less competitiveness on the part of women and treat them differently accordingly.

evidence that nurture drives at least part of the observed gap. Similarly, our findings suggest policies that seek to reduce the gender gap in competitive preferences should consider the importance of shifting social norms.

In a broader sense, our paper comments on the importance of subject pool differences in evaluating the external validity of experimental results. Previous work on experimental methodology has highlighted the importance of subject characteristics but tends to focus on concerns regarding using students who are convenient but may not adequately represent the population (Abbink and Rockenbach, 2006; List, 2003). For similar reasons, the majority of experiments are run using subject pools that are predominantly Caucasian; our work highlights that some results using such subjects may not generalize to other cultures. Rather than a problem, we see this observation as an opportunity. Our work highlights the importance of multiple identities interacting and moderating the effect of group specific norms. Although there is substantial heterogeneity in preferences across populations, it is difficult to identify the root of these differences as preferences form endogenously over lengthy periods. Our work illuminates the value of harnessing cultural variation to deepen our understanding of the interaction of multiple identities and its effect on preferences and their resulting behavior.

2.2 Literature Review

Our work enriches an existing literature focused on understanding the persistence of the gender gap in earnings and promotion. Historically, researchers have argued for the importance of statistical discrimination driven by beliefs that men are innately better suited for certain tasks or that women are more likely to leave their job to raise a family (Arrow, 1973; Polachek, 1981). Looking at a separate aspect of discrimination, Goldin and Rouse (2000) find that introducing blind auditions increases female representation in orchestras suggesting the existence of taste-

based discrimination. Instead, we join others in investigating behavioral explanations for the observed gender gap in promotion and wages. Examples of gender differences in behavior include that women are less likely to negotiate for higher salaries (Babcock and Laschever, 2009) and are more likely to complete low promotability tasks (Babcock et al., 2017). These differences provide an alternative explanation for the persistence of the gender gap even in the presence of greater overall gender equality.

Niederle and Vesterlund (2007) investigate an alternative behavioral explanation in the form of competitive preferences. In their design, participants first complete an addition task under a piece rate payment scheme in which payment is received independent of the performance of others. In a second task, participants complete the same addition task under a competitive payment scheme in which they are only paid if they are the top performer in a group of four participants. After experiencing both payment schemes, participants select which payment scheme they would prefer to be applied to a third performance. Although women perform equally well as men, male participants were close to twice as likely to pick the competitive pay scheme in the third task. Overconfidence among males drives part of these results, but differences in beliefs alone cannot fully explain the large gap in competition entry. The authors suggest that the difference in competitive preferences explains the lack of female representation among top earning positions.

Others have replicated this result across various tasks including mazes (Booth and Nolen, 2012), word search (Cárdenas et al. 2012), and ball throwing (Gneezy et al. 2009; Anderson et al. 2013). Similarly, the gender gap in competitive preferences has been found in a wide range of environments including field settings ranging from tribes generally outside the developed world (Gneezy et al. 2009; Anderson et al. 2013) to school children of various ages and in different countries (Booth and Nolen 2012; Cárdenas et al. 2012).

In a follow-up paper, Niederle et al. (2013) find that introducing an affirmative action policy that guarantees a winning position to a member of each gender closes the gap in competitive preferences at little cost to efficiency. Others find that introducing a team component to competition (Healy and Pate, 2011; Dargnies 2012) or simply providing feedback (Wozniak et al., 2014; Ertac and Szentes, 2011) can shrink the gender gap in competitive preferences. Taken together, this research suggests that carefully designed institutions can shrink the existing gap in competitive preferences.

Instead, we join others in seeking to understand the mechanism behind the gender gap in competitive preferences. Some researchers suggest men are more competitive due to the pressures of reproductive success: men who compete successfully are more likely to attract a high quality mate (Daly and Wilson, 1983), an argument that favors nature as the root of the competition gap and the notion that men are innately better at competition (Baron-Cohen and Benenson, 2003). However, Akerloff and Kranton's (2000) theoretical work on the influence of identity suggests that group specific norms may be an important driver of behavior. If this is true, then gender norms suggesting that women should be cooperative and men should be competitive could drive the observed gender gap in competitive preferences. In line with this possibility, Akerloff and Kranton's (2000) theoretical work is complemented by anecdotal evidence of such a tension for female lawyers whose identity as a lawyer encourages them to be competitive but whose identity as a woman encourages the opposite.

Investigating the role of nature and nurture in the development of competitive preferences, Anderson et al. (2013) and Cárdenas et al. (2012) find that there is no gap in competitive preferences among young children. These authors argue that puberty marks an important milestone in the presence of gender differences in preferences. Yet it is unclear whether nature or nurture (or

a combination of both) explain this result; one possibility is that the observed gender differences in preferences are the results of different hormonal fluctuations that do not occur until puberty. In support of this possibility, Wozniak et al. (2014) and Buser (2012) find that the menstrual cycle and the associated cascade of hormones can predict the presence of a gender gap in competitive preferences. Taken together, these works suggest that nature plays a role in the formation of competitive preferences. However, it is similarly plausible that the mechanism behind the role puberty plays in this gap is through the recognition of gender norms that comes with sexual maturity. In favor of this explanation, Booth and Nolen (2012) compare teenage females who attend single sex and co-ed institutions and find that girls who attend single sex institutions have competitive preferences similar to the men rather than to the women who attend co-ed schools. One possible explanation for this result is that the difference in learning environment affected the saliency of gender norms and therefore their effect on behavior. In line with this hypothesis, Almås et al. (2015) documents the standard gap in competitive preferences but finds that individuals in the top income quartile drive the gap. The authors argue that scarcity and may moderate the effect of norms in the exhibition of competitive preferences. Similarly, Cassar et al. (2016) find the standard gap when the prize of the competition is monetary but no such gap when the prize is a school voucher. They find that women are willing to compete, but they reserve their competitive tendencies for the success of their offspring; similar to Almås et al. (2015), these authors find that women are willing to compete when they deem the need great enough. Given these results, it is unlikely nature alone can explain the formation of competitive preferences as females' preference for competition is environment dependent.

Consistent with the previous results on the importance of environment, Gneezy et al. (2009) were the first researchers to document the existence of a culture exhibiting no gender gap in

competition. The authors find the standard gap in competitive preferences among the patrilineal Masai but find that women were more likely to compete than men among the matrilineal Khasi. More recently, Dariel et al. (2017) find no gender gap in competitive preferences among Emiratis and argue that their results are driven by cultural transition away from the traditional patriarchal structure in the United Arab Emirates. Given it unlikely that these populations faced a different set of evolutionary pressures, the authors find support in favor of nurture playing a role in the development of competitive norms.[2]

We contribute to this body of work by documenting the first adult population in the United States to exhibit no gender gap in competitive preferences. Directly replicating Niederle and Vesterlund's (2007) original procedure with two novel populations, we find that Hispanic students exhibit the standard gap in competitive preferences while no such gap exists among black students. Our finding documents a population, which differs solely in culture, exhibiting no gender gap in competitive preferences. Although others have documented populations in which there is no gender gap in preferences, we are the first, to our knowledge, to provide clear evidence that culture and norms must explain at least part of the observed gender gap in competitive preferences. Our work relates to Gneezy et al. (2009) and Dariel et al. (2017) with one important distinction: beyond cultural variation, the populations they studied experienced different institutions that may confound the effects of culture on competitive preferences. Our results illuminate the importance of identity in the form of social norms in determining the expression of competitive preferences. Additionally, our work highlights the importance of multiple identities in moderating the effect of norms on the formation of preferences and their resulting behavior.

[2] Gong and Yang (2012), Liu and Huang (2013), and Falk and Hermle (2018) find that another documented gender difference, risk preference, is context dependent as well.

2.3 Experimental Design

Our design closely follows the experimental design first conducted by Niederle and Vesterlund (2007) and was conducted using the Zurich Toolbox for Ready Made Economic Experiments (ZTREE; Fischbacher, 2007). We assign participants to groups containing two men and two women. Although we do not explicitly mention gender, participants see the other members of their group and therefore its gender composition as well. As in the original study, we ask participants to add up random sets of five 2-digit numbers for a total of five minutes under different payment schemes. We use this task to avoid gender differences in ability and to ensure comparability between our results and those of the original paper. After a participant submits an answer, a new set of randomly generated 2-digit numbers appears as well as feedback on whether their previous answer was correct. Participants receive no feedback about the performance of their fellow group members at any point, but receive absolute feedback on their performance following the conclusion of each task. Following Niederle and Vesterlund (2007), we provide scrap paper but no access to a calculator. Participants completed this task under a number of payment schemes in the order described below:

Task 1 (Piece rate pay scheme): This is the baseline used to measure individual's ability under a non-competitive piece rate scheme and ensure no gender differences in ability. If this task is selected for payment, subjects receive $0.50 for each correct answer and their earnings are independent of the performance of the other members of their group.

Task 2 (Competitive pay scheme): This task is identical to task 1 with the important exception that in each group the individual who answers the most questions correctly receives $2.00 per correct answer while the other members of the group receive nothing. This payment is designed so that for a given number of correct answers, a risk-neutral participant with a 25%

chance of winning should be indifferent between a competitive and non-competitive payment scheme. This task is included to investigate whether simply being in a competitive environment creates a gender difference in performance.

Task 3 (Choice): Before the addition task, participants are asked to select which payment scheme they would like to be applied to their performance in the case that this task is selected for payment. Thus, participants are deciding whether they want to receive $0.50 per correct answer or $2.00 per correct answer conditional on being the top performer on their team. If a participant picks the competitive scheme, their performance is compared to their group member's performance on the previous task. This task measures individual's preferences for competition.

Task 4 (Choice for previous performance): Unlike the previous tasks, participants do not have to perform the addition task. Instead, they must decide whether they would like their task 1 performance to have a piece rate or competitive payment scheme applied. Thus, participants are choosing whether they want to receive $0.50 per correct answer from task 1 or $2.00 per correct answer in the case where they were the top performer in their group in task 1. This task is necessary to identify whether having to perform in a competitive atmosphere has an impact on the gender gap in competitive preferences.

After completing all four tasks, we introduce an incentivized belief elicitation to investigate how confidence relates to competitive preferences. We ask participants to guess where they ranked among their team in task 1 and 2, and they receive $1 per each correct guess. All of the procedures listed above are identical to those used in Niederle and Vesterlund (2007). We depart from their design through the inclusion of an incentivized risk elicitation measure as well as a survey designed to identify potential cultural correlates of competitive preferences.

Our incentivized risk elicitation measure is adapted from Eckel and Grossman (2008), as modified in Dave et al. (2010) and then scaled down to be consistent with the payments from the other tasks. The measure works as follows: we ask participants to select from among six 50-50 gambles. Each successive gamble increases in expected return as well as variance. If this task is selected for payment, the computer generates a random number for each participant, and if the number generated was above 0.5 the participant receives payment A for the gamble they selected and payment B otherwise. The choice of gamble indicates a degree of risk tolerance, from none (Gamble 1) to risk-seeking (Gamble 6).

Our survey is included in the appendix. It was designed to measure cultural differences that may correlate and perhaps explain differences in the gender gap in competitive preferences across different individuals and groups. Items in the survey pertain to individual's socioeconomic status, family structure, ethnicity, self-efficacy and beliefs about proper gender roles.

We ran the experiment at two locations: The Economics Research Laboratory (ERL) at Texas A&M University and the Prairie View Research Laboratory (PVRL) at Prairie View A&M University. We recruited subjects at both locations using the Online Rescruiting System for Economics Experiments (ORSEE; Griener, 2015). We ran our Hispanic participant sessions at the ERL and the vast majority of our black participant sessions at PVRL. These locations are ideal for investigating the role culture may play in competitive preferences by providing access to subjects from two populations that are understudied in the competition literature. Although these locations are ideal in providing access to unique subjects, there are significant differences across the two locations. Texas A&M is a large state school hosting over 53,000 undergraduate students whereas Prairie View A&M University is significantly smaller hosting only 9,125 undergraduate students. The acceptance rates and cost of attendance are similar at each university, but Texas A&M requires

significantly higher average SAT scores (1250 vs. 993). Additionally the demographics of each school are significantly different: 23.6% of Texas A&M undergraduate students are Hispanic while the same is only true of 9.4% of undergraduate students at Prairie View A&M University. On the other hand, 84.7% of undergraduate students are African American at Prairie View A&M University (making it a Historically Black University) while the same is only true for 3% of the undergraduate students at Texas A&M University. Lastly, Texas A&M has a roughly equal split of male and female undergraduates (47.8% female) while Prairie View A&M has significantly more female undergraduates (61% female). In an effort to test for selection concerns, we ran two sessions of black participants at Texas A&M (n=20) and compare the behavior of black students across campuses and discuss why the differences across campuses are unlikely to explain our findings in the results section.

In total, we have 96 Hispanic and 128 black subjects. On average, subjects earn $20.54 and there are no discernible differences in earnings across gender or ethnicity. Sessions included 1 to 4 groups from a single university and ethnicity. Sessions took less than an hour and participants received a $5 show-up fee plus $7 for completing the experiment plus whatever they earned in the selected task. Following Niederle and Vesterlund (2007), we pay for a single task to avoid decisions in one task affecting decisions in another task to hedge over earnings. We ran the first wave of sessions in April 2017 as part of an undergraduate research course. During this wave, five sessions were run at Texas A&M and four sessions were run at Prairie View A&M University producing 56 Hispanic and 64 black subjects respectively. We carried out a second wave of sessions during the spring of 2018 including an additional six sessions at Texas A&M University producing 40 Hispanic and 20 black subjects. These black subjects serve as an important control

for possible concerns over selection as a driver of the primary results. We also ran an additional five sessions at Prairie View A&M University producing an additional 64 black subjects.

2.4 Theory and Hypotheses

We base our hypotheses on the theoretical work of Akerloff and Kranton (2000). In their model, individuals receive increased utility from being members of groups or put another way by having identity. Akerloff and Kranton (2000) argue that the magnitude benefit is determined by how much an individual values a particular identity and how much the group accepts the individual. As a result, individuals have a desire to conform to the norms of a group to increase their acceptance by other group members. Taken together, the intuition provided by the model in forming our hypotheses is that individuals adjust their behavior to follow the social prescriptions of the communities to which they belong.

In the experiment, the choice of payment scheme is between a low rate, which is guaranteed independent of the performance of others, or a high rate, which is only paid in the case that one is a top performer. In the absence of identity considerations, ability, confidence (beliefs about one's own ability), and one's taste for risk should drive competition entry as these determine the expected utility of each payment scheme. However, identity changes this decision by suggesting what the individual ought to do. In this sense, group identity can lead to additional costs to entering a competitive environment. Sociologists have identified the tendency for Western culture to exhibit such norms regarding gender: competitiveness is good for men but bad for women (Lorber and Farrell, 1991; Rose and Rudolph, 2006). Since females exhibiting aggressive or competitive behavior may be acting contrary to their identities prescribed behavior, there may be additional costs applied to females entering competition (Akerloff and Kranton 2000). Under these conditions, the model would predict that a woman would be less likely to enter competition than a

man of similar ability, beliefs, and risk preferences. At an aggregate level, this would produce the observed gender gap in competitive preferences.

Beyond this, the model also can accommodate the absence of a gap in certain environments. Matrilineal societies (Gneezy et al. 2009) have female-led households, and this may create a norm supporting female competitiveness. Single-sex schools (Booth and Nolen 2012) reduce the salience of gender norms, allowing women to avoid the identity-based utility loss from entering competition. Moreover, low socioeconomic status (Almås et al. 2015) implies scarcity that crowds out the cost of norm violation especially in the case of mothers.

The model similarly explains the success of institutions in decreasing the gap. The introduction of affirmative action policies (Niederle et al. 2013) replaces the standard norm by implying women should compete. Team competition (Healy and Pate, 2011; Dargnies 2012) adds a cooperative element, a trait that women are expected to exhibit. Lastly, because women are expected to ensure the survival of offspring, introducing an incentive scheme directly related to the success of one's offspring shifts female competitiveness to being viewed positively (Cassar et al. 2016).

Although a number of differences exist between Hispanic and black culture, we focus on the differences in gender norms across populations. Sociologists find suggestive survey evidence that Hispanic individuals hold less egalitarian gender norms than white individuals (Harris and Firestone, 1995; Kane, 2000). Within our theoretical framework, this suggests potentially larger differences in gender competitiveness in our Hispanic sample relative to the original study. On the other hand, the black population contains a large number of single-mother and female-headed households (U.S. Census Bureau, 2018). In the case of single motherhood, the mother is the sole provider for her children, a circumstance that encourages female competitiveness. We hypothesize

that this in turn produces daughters who learn that being competitive is good. In line with this hypothesis, sociologists find that black individuals tend to hold more positive beliefs about maternal employment (Blee and Tickamyer, 1995) and tend to show greater recognition of gender inequality (Hunter and Sellers, 1998). Given these assumed cultural norms regarding competitiveness, the model predicts a difference across our subject pools.

Hypothesis 1 (Hispanic Competitiveness): We expect to observe a gender gap in competitive preferences in the Hispanic population. Controlling for differences in ability, Hispanic men will be significantly more likely to opt for a competitive pay scheme relative to their female counterparts.

Hypothesis 1a (Hispanic Competitiveness relative to Original Study): We expect to observe a gender gap in competitive preferences in the Hispanic population that is larger than the original study which was conducted with a predominantly white sample.

Hypothesis 2 (Black Competitiveness): We expect to observe no gender gap in competitive preferences in the black population. Controlling for differences in ability, black men and women will be equally likely to opt for a competitive pay scheme.

2.5 Results

In the following section, we discuss the results of our experiment. We begin by comparing the choice of payment scheme in task 3 across gender and ethnicity followed by investigating how differences in performance and beliefs about performance may explain our results. We repeat the same analysis for the choice of payment scheme in task 4 and conclude by discussing selection concerns and the importance of culture and risk aversion as measured by our survey.

2.5.1 Tournament Entry

Turning first to the main measure of interest, the rate of entry into the tournament, when the data is pooled, men opt for tournament payment about 52% of the time while the same is only true for 43% of the women. This difference is significant (Fisher's exact test, $p<0.05$) but reflects a much smaller gap than reported in the original Niederle and Vesterlund (2007) paper (9% vs. 35%).

Figure 2.1: Proportion of Subjects Selecting Competitive Payment Scheme in Task 3

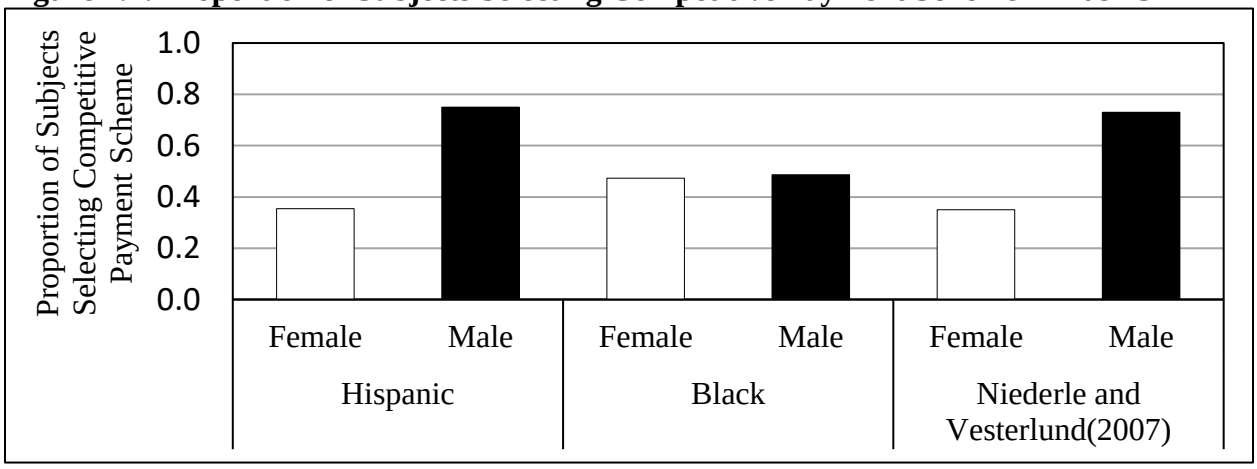


Figure 2.1 presents the proportion of individuals choosing the competitive payment scheme in each sample split by gender. When we divide the entry decisions by ethnicity, we find that the gender gap in competition entry within the Hispanic sample is significantly different from the one found within the black sample. Hispanic men opt for a competitive pay scheme more than twice as often as Hispanic women (75% and 35% respectively). This gap in competitive preferences is statistically significant (Fisher's exact, $p<0.001$), but not significantly different from the original results. On the other hand, our black sample exhibits no significant gender gap with men and women entering competition at similar rates (49% and 47% respectively, Fisher's Exact, $p=1.00$).

The aggregate data support our hypothesis that culture plays a role in the development of competitive preferences. However, it is possible that performance differences rather than differences in competitive preferences explains the difference in competition entry across samples.

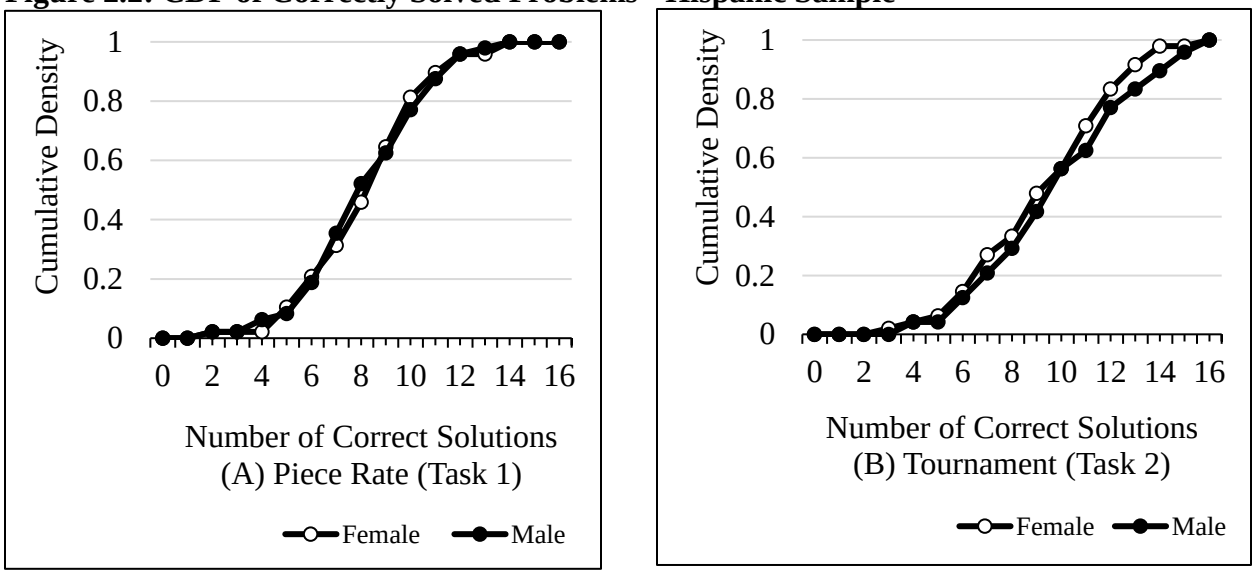
2.5.2 Performance Differences

We detect no significant gender difference in performance on the addition task. In the piece rate payment (task 1), men correctly answer an average of 6.91 questions while women correctly answer 7.30 questions; the difference in distributions of correct answers is not statistically significant (Mann-Whitney, $p=0.22$). Looking towards task 2 (tournament payment), men correctly answer 8.50 questions relative to 8.59 questions for women. Although the improvement across tasks is significant for both men and women (two-sided t-test, $p<0.001$ for both), neither improvement nor performance in task 2 are significantly different across gender (Mann-Whitney, $p=0.34$ and $p=0.90$ respectively). Among the 61 groups created during the experiment, women were the top performer in their group in second task 31 times, while men were the top performer in their group 29 times. We conclude that there exists no difference in performance across gender. However, it is possible that this approach masks heterogeneity in performance within our sampled populations.

Figure 2.2 presents the cumulative distribution of performance on the first and second task for the Hispanic sample in the left and right panel respectively. For every performance level, the graph displays the number of Hispanic men and women who did as well or worse on the addition task. In task 1 (piece rate payment), Hispanic men solved an average of 8.54 problems compared to 8.58 problems for the women. This difference is not significant (Mann-Whitney, $p=0.96$). Turning to the task 2, we continue to detect no significant performance differences in the Hispanic sample. Hispanic men correctly answer an average of 10.23 problems while Hispanic women

correctly answer 9.67 questions, (Mann-Whitney, p=0.42). Similar to our aggregate analysis, the improvement in performance is significant for Hispanic men and women (two-sided t-test, p<0.01 for both) but not significantly different across gender (Mann-Whitney, p=0.15). Among the 24 Hispanic groups, men and women were the top performer in their group 12 times each. Taken together, we conclude that performance differences cannot explain any observed difference in competitive preferences within the Hispanic population.

Figure 2.2: CDF of Correctly Solved Problems - Hispanic Sample



Repeating the same analysis for the black sample (figure 2.3), we find that black women slightly outperform men on task 1 answering 6.45 questions correctly on average compared to 5.85 questions for the men (Mann-Whitney, p=0.06). However, looking at performance under a competitive pay scheme (task 2), the marginal difference is eliminated; men answer an average of 7.38 problems compared to an average of 7.89 for the females (Mann-Whitney, p=0.37). Although the improvement across tasks is significant for both men and women (two-sided t-test, p<0.001

for both), the difference in improvement is not (Mann-Whitney, p=0.95). Among the 37 black groups, women were the top performer in their group 20 times. Again, we conclude that performance differences cannot explain any observed gap in competitive preferences.

Figure 2.3: CDF of Correctly Solved Problems- Black Sample

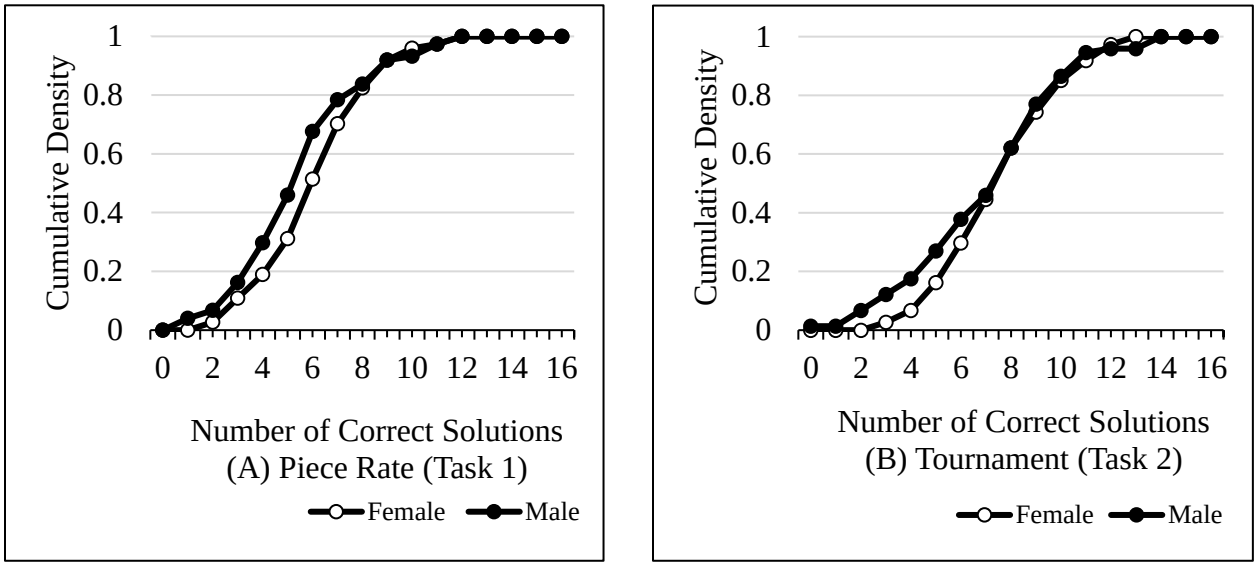
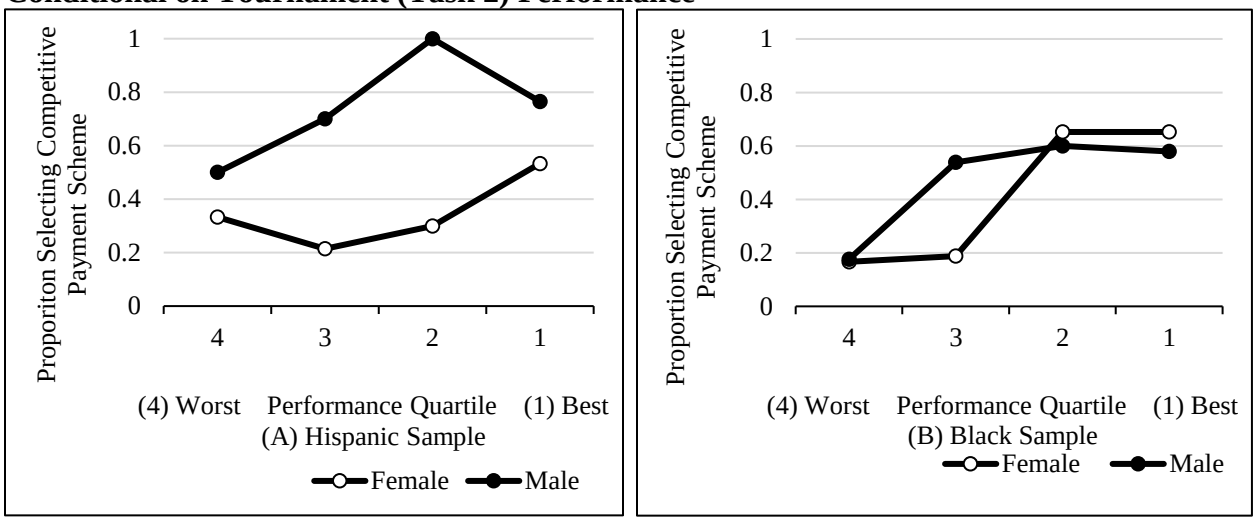


Figure 2.4 presents the competition entry decisions of Hispanic (left panel) and black (right panel) subjects conditional on their previous performance under a competitive pay scheme (task 2). The left panel illustrates that for every performance quartile Hispanic men are more likely to opt for a competitive pay scheme than their female counterparts. Although more Hispanic women enter as their performance improves, Hispanic women from the top quartile are as likely to enter as Hispanic men from the worst quartile. We conclude that Hispanic men opt for competition at a higher rate regardless of actual performance. Turning to the second panel, we see that among our black sample the gender gap in competition entry that is found in panel A only occurs for those in the third quartile of performance. Although the difference in entry is marginally significant among

those in third quartile of performance (Fisher's exact, p=0.06), in every other quartile black women are equally or more likely to opt for competition than their male counterparts.

Figure 2.4: Proportion of Subjects Selecting Competitive Payment Scheme for Task 3 Conditional on Tournament (Task 2) Performance



To investigate the role that performance plays in competition entry and the observed gender gap, we estimate a probit model of competition entry conditioned on gender, ethnicity, performance in the second task and the amount of improvement between task 1 and task 2. We present the results of this regression in table 2.1. Consistent with the earlier analysis, being a female has a large negative impact on the probability of entering competition. However, this is only true of Hispanic females as the positive effect of being black conditional on being a female is almost as large as the coefficient on the female dummy. Overall, we conclude that performance differences cannot explain the observed pattern of competition entry.

Table 2.1: Probit of Competition Entry in Task 3

	(1)
Female	-1.024 (0.264)***
Black	-0.452 (0.246)*
Black x Female	0.939 (0.320)***
Tournament (Task 2) Performance	0.098 (0.039)***
Tournament (Task 2) - Piece Rate (Task 1)	0.017 (0.048)
N	244

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is task 3 choice of compensation scheme (1-tournament 0-piece rate)

2.5.3 Beliefs about Tournament (Task 2) Performance

Although performance differences do not appear to explain the gap in competitive preferences, it is possible that a discontinuity between performance and beliefs about performance explains our results. To investigate this concern, figure 2.5 presents the frequency of guessed rank in task 2 for the Hispanic (left panel) and black (right panel) subjects. Looking at the left panel, we can see Hispanic men exhibit a significantly higher level of confidence and are over twice as likely to believe they were the best performer in their group (two-sided t-test, $p<0.01$). Given that we detected no difference in performance on the task, Hispanic males must be overconfident relative to their female counterparts. In line with this, Hispanic males are more likely to incorrectly rank themselves, and, in every case of incorrect ranking, Hispanic males ranked themselves higher than their true ranking within their group (table 2.2).

Figure 2.5: Distribution of Guessed Tournament (Task 2) Rank

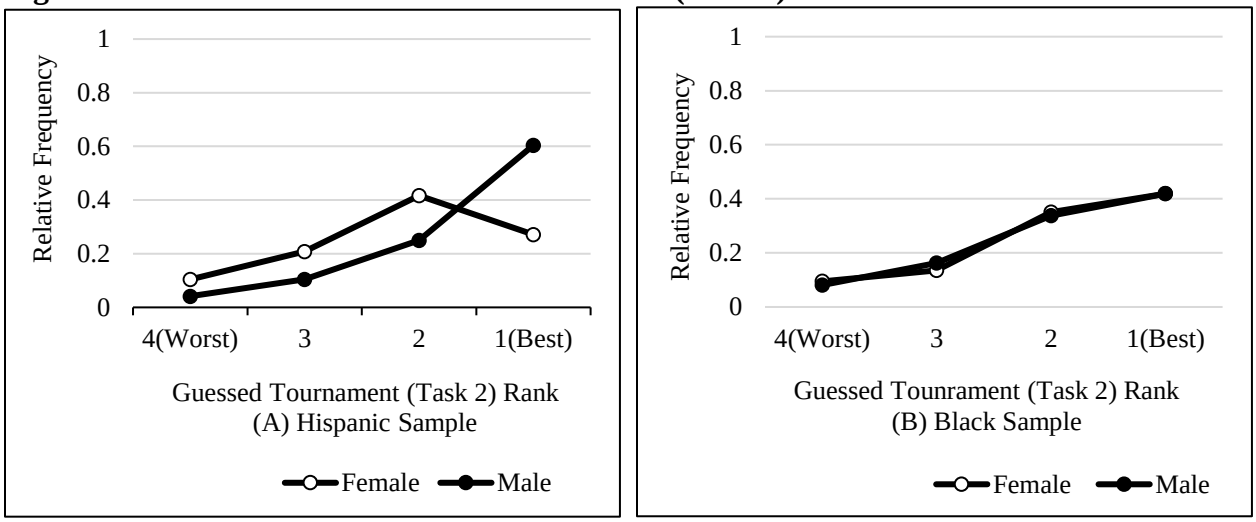


Table 2.2: Accuracy of Guessed Tournament Rank (Task 2)

	Hispanic				Black			
	Female		Male		Female		Male	
	Guessed	Incorrect	Guessed	Incorrect	Guessed	Incorrect	Guessed	Incorrect
1(Best)	13	3	29	15	31	15	31	22
2	20	11	12	7	26	16	25	16
3	10	4	5	1	10	3	12	12
4(Worst)	5	0	2	0	7	4	6	4
Total	48	**18 (37%)**	48	**26 (54%)**	74	**38 (51%)**	74	**54(73%)**
Overconfident		**15 (83%)**		**26 (100%)**		**27(71%)**		**42(78%)**

Turning to the black sample, we see the distribution of beliefs are almost identical for every belief quartile. Black males and females are equally likely to guess they are the best performers in their group (two-sided t-test, p=1.00). Similar to the Hispanic population, males tend to incorrectly rank themselves more often. However, the probability of over-ranking conditional on being incorrect is similar for black men and women. To confirm the intuition provided by the figures,

we conduct an ordered probit estimating guessed rank based on gender and actual performance (table 2.3).

Table 2.3: Ordered Probit of Guessed Tournament Rank (Task 2)

	(1)
Female	0.853 (0.165)***
Black	-0.214 (0.255)
Black x Female	-0.686 (0.272)**
Tournament (Task 2) Performance	-0.296 (0.043)***
Tournament (Task 2) - Piece Rate (Task 1)	-0.023 (0.044)
N	244

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is guessed tournament (task 2) rank (1-best, 4-worst)

The results of table 2.3 confirm the intuition provided by table 2.2 and figure 2.5. After controlling for performance, men are still significantly more confident relative to their female counterparts. Similar to entry decisions, this effect is much larger for Hispanic sample. We conclude that males are overconfident in their beliefs, but the gap in confidence is much larger for the Hispanic sample. We investigate the role overconfidence plays in explaining the observed gap in competitive preferences by repeating the earlier probit analysis with the additional control of participant beliefs.

The repeated probit analysis (table 2.4) shows that beliefs have a strong positive relationship with competition entry and that performance impacts competition entry primarily through beliefs.[3] Overconfidence explains part of the gap in entry, but beliefs alone cannot explain

[3] Recall that a lower rank means higher confidence, as a rank of 1 is a belief in being the best performer, while a rank of 4 is a belief in being the worst performer. Thus, the negative coefficient captures a positive relationship between confidence and entry.

the pattern of competition entry in our sample. Overall, we conclude that the observed pattern of behavior is robust to alternative explanations involving beliefs and performance differences as these alone cannot explain the presence or lack of a gender gap in tournament entry in our samples.

Table 2.4: Repeated Probit Analysis of Task 3 Competition Entry with Beliefs about Relative Performance

	(1)	(2)
Female	-1.024 (0.264)***	-0.837 (0.273)***
Black	-0.452 (0.246)*	-0.588 (0.230)**
Black x Female	0.939 (0.320)***	0.792 (0.316)**
Tournament (Task 2) Performance	0.098 (0.039)***	0.003 (0.045)
Tournament (Task2)-Piece Rate (Task1)	0.017 (0.048)	0.012 (0.049)
Guessed Tournament Rank		-0.517 (0.138)***
N	244	244

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is task 3 choice of compensation scheme (1-tournament 0-piece rate)

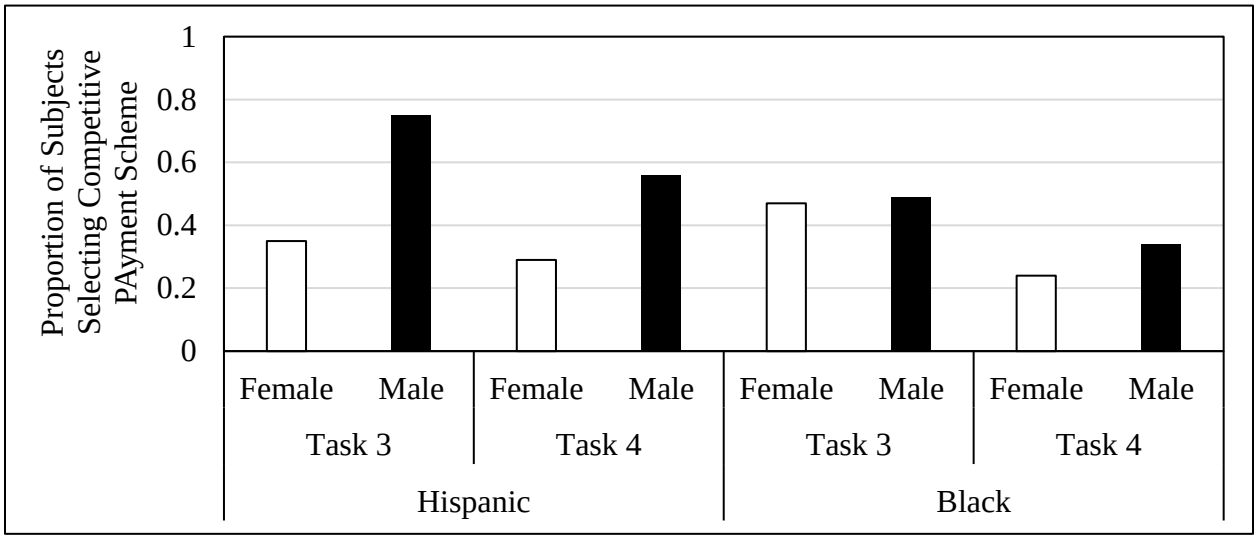
2.5.4 Choice for Past Performance (Task 4) Analysis

Niederle and Vesterlund (2007) highlight an alternative explanation that their results may reflect preferences over having to perform in competitive environment instead of preferences towards competitive incentives. To test for this concern, task 4 is introduced. In this task, participants did not have to perform the real effort task. Instead, they simply chose whether they would prefer to have a tournament or piece rate payment applied to their first task performance.

Figure 2.6 presents a comparison of entry by gender and ethnicity for both task 3 and task 4. Although the gap in entry for task 4 is smaller for the Hispanic sample, the gap is still large and significant (Fisher's exact, $p<0.05$). The gap in the black sample is bigger in task 4 relative to task

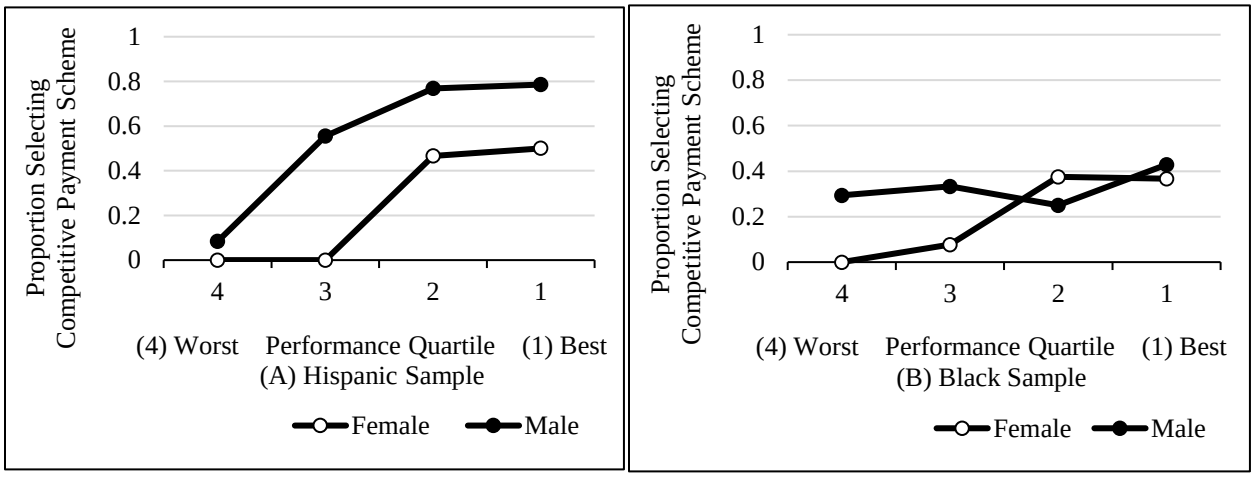
3, but the gap in entry is still not significant (Fisher's exact, p=0.28). On the surface, we conclude that apprehension for competition does not explain the gender gap in entry. However, similar to task 3, it is still possible that performance differences or beliefs explain the presence or lack of a significant gender gap in our samples.

Figure 2.6: Comparison of Proportion of Subjects Selecting Competitive Payment Scheme in Task 3 and Task 4



We delve into the relationship between payment choice in task 4 and performance in task 1 in figure 2.7. Looking at the Hispanic sample (left panel), both genders exhibits a strong positive relationship between performance and opting for a competitive pay scheme in task 4. There is still clear separation in entry illustrating that Hispanic men opt for a competitive scheme more often than Hispanic women at every performance quartile. On the other hand, the black sample appears to exhibit no gap in entry among those who should enter (the top two quartiles), but males who perform in the lower quartiles appear to over-enter similar to Hispanic males. Performance seems to have little effect on competition entry for black males, as they are equally likely to enter if they are in the lowest or highest quartile.

Figure 2.7: Proportion of Participants Selecting Competitive Payment Scheme in Task 4 Conditional on Task 1 Performance



Repeating our earlier probit analysis, table 2.5 suggests apprehension to having to perform in a competitive environment cannot explain the observed gender gap in entry as females are still less likely to opt for a tournament payment scheme after controlling for performance. Although the gender effect is smaller for the black sample, we can no longer claim that there is no gap in entry for the black population. However, in tangent with the task 3 results, this would imply black women are eager rather than apprehensive to compete. Before concluding our task 4 analysis, we consider the role that beliefs play in tournament entry in task 4.

Table 2.5: Probit of Competition Entry in Task 4

	(1)
Female	-0.796 (0.279)***
Black	-0.127 (0.262)
Black x Female	0.379 (0.358)
Piece Rate (Task 1) Performance	0.180 (0.039)***
N	244

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is task-4 choice of compensation scheme (1-tournament 0-piece rate)

Figure 2.8 presents the distribution of guessed ranks in task 1 for the Hispanic (left panel) and black (right panel) samples. Similar to task 3, Hispanic men ranked themselves higher than Hispanic women with close to half the men believing they were the best performer in their group. Although not as large as the difference for the Hispanic population, black men are also more likely to believe they were the top performer in their group. However, the difference in average ranking is only significant for the Hispanic population.

Figure 2.8: Distribution of Guessed Piece Rate (Task 1) Rank

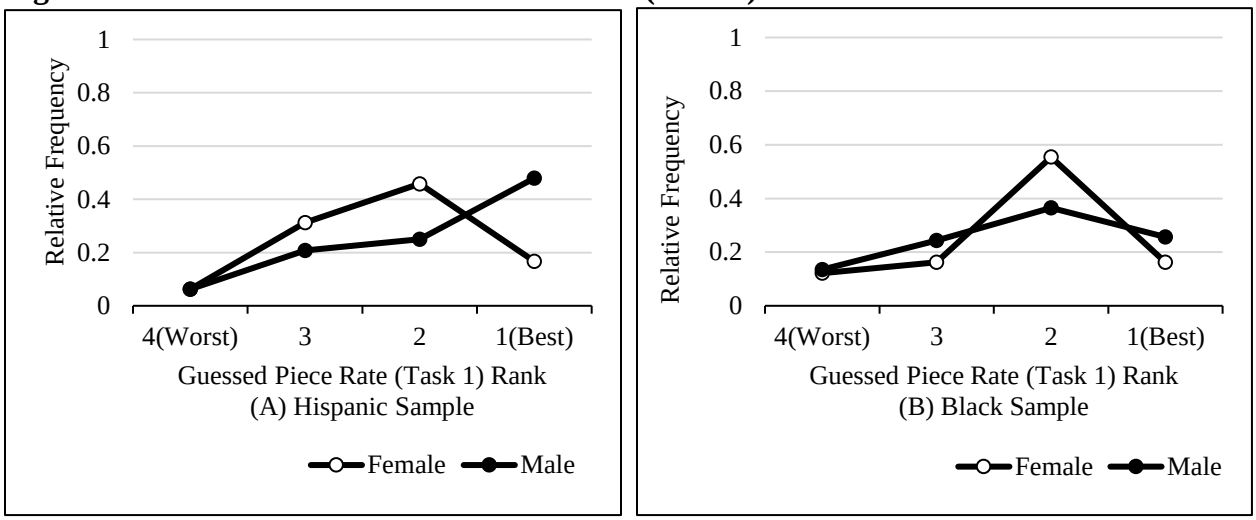


Table 2.6 presents an ordered probit model predicting guessed piece rate rank. In line with the task 3 results, females rank themselves significantly lower but the effect is much smaller for black sample. To see how the difference in beliefs predicts competition entry we repeat the probit analysis with the added control of guessed piece rate rank in table 2.7. Again, we find that overconfidence explains part of the gender gap in entry, but females are still significantly less likely to submit their past performance for a competitive payment even after controlling for beliefs. Taken together, the patterns of results from task 4 are similar to those from task 3 leading us to

conclude that apprehension to competing cannot explain the observed difference in competitive preferences.

Table 2.6: Ordered Probit of Guessed Piece Rate Rank (Task 1)

	(1)
Female	0.723 (0.233)***
Black	-0.282 (0.229)
Black x Female	-0.484 (0.296)
Piece Rate (Task 1) Performance	-0.342 (0.038)***
N	244

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is guessed piece rate (task 1) rank (1-best, 4-worst)

Table 2.7: Repeated Probit Analysis of Task 4 Competition Entry with Beliefs about Relative Performance

	(1)	(2)
Female	-0.796 (0.279)***	-0.600 (0.295)**
Black	-0.127 (0.262)*	-0.238 (0.260)
Black x Female	0.379 (0.358)	0.240 (0.372)
Piece Rate (Task 1) Performance	0.180 (0.039)***	0.064 (0.047)
Guessed Piece Rate Rank		-0.543 (0.127)***
N	244	244

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is task 4 choice of compensation scheme (1-tournament 0-piece rate)

2.5.5 Selection

A final alternative explanation for our results is school selection. Texas A&M University is more competitive than Prairie View A&M University and thus student's choice of university may reflect differences in preferences for competition. Yet, logically if women tend to avoid

competitive environments, we would expect selection to produce an effect opposing our results: competitive women would select into Texas A&M while the effect on males should be limited as men are "innately" competitive. Thus, we would predict no gap in Preferences at Texas A&M and the standard gap at Prairie View A&M.

To investigate the possibility of selection driving our results, we ran two sessions of black students at Texas A&M. table 2.8 presents a comparison of our black subjects from each university along dimensions that may affect competitive preferences. The only notable difference is that the gap in competitive preferences is larger among Texas A&M black students. Of note, this gap is much smaller than that found in the Hispanic sample (10% vs. 35%). This difference may be predicted from our model of behavior. As Texas A&M tends to be more competitive and a predominantly white college, the associated social norms may be more salient for black females at Texas A&M relative to their Prairie View counterparts causing them to act behave less competitively.

Table 2.8: Comparison across Campuses

	Texas A&M University- Hispanic (N=96)	Texas A&M University-Black (N=20)	Prairie View A&M University (N=128)
Number of Problems Solved (Average of Task 1 + 2)	9.26	7.30	6.84
Women	9.13	7.20	7.18
Men	9.39	7.40	6.49
Tournament Entry	0.55	0.45	0.48
Women:	0.35	0.40	0.48
Men:	0.75	0.50	0.48
Guessed Tournament Rank (1=Best) (4=Worst)	1.86	2.05	1.88
Women:	2.15	2.00	1.89
Men:	1.58	2.10	1.88

To test whether selection can explain our results, we repeat all of our regression analysis including dummies for black students at Texas A&M in the appendix. The magnitudes of coefficients change in some circumstances but the coefficients on the A&M black dummies never attain significance and their inclusion does not change the qualitative conclusions: Hispanic subjects exhibit the standard gender gap in competition while no such gap exists in the black sample. We conclude that it is unlikely that selection at the university level explains our pattern of results.

2.5.6 Survey

After the experimental conditions, subjects completed a short survey including an incentivized and self-reported risk measure[4], questions about ethnicity, and measures of self-image. Because the competitive payment scheme introduces risk into payment, it is possible that our results are a reflection of risk preferences rather than competitive preferences. Consistent with the original paper (Eckel and Grossman, 2008), we find the females in both samples are more risk averse than their male counterparts in an incentivized setting, but the difference in risk preferences is only significant for the black sample (two-sided t-test, $p=0.14$ for Hispanics, $p<0.01$ for blacks). In a second measure taken from the National Longitudinal Survey of Youth, women in both samples are significantly less willing to take risks (two-sided t-test, $p<0.001$ for both samples). We repeat the probit analysis of competition entry in task 3 adding both risk measures in table 9. Both risk measures predict competition entry and explain part of the gap but the female and female black interaction are still significant. We conclude that differences in risk preferences contribute to the observed gap but cannot fully explain it.

[4] The self-reported risk measure was taken from the National Longitudinal Survey of Youth and ask participants to rank how willing they are to take risks on a scale from 1 to 10 with 1 representing being "completely unwilling to take risks" 10 representing being "very willing to take risks."

Table 2.9: Repeated Probit Analysis of Task 3 Competition Entry with Beliefs about Relative Performance and Measures of Risk Aversion

	(1)	(2)	(3)
Female	-1.024 (0.264)***	-0.837 (0.273)***	-0.631 (0.283)**
Black	-0.452 (0.246)*	-0.588 (0.230)**	-0.540 (0.228)**
Black x Female	0.939 (0.320)***	0.792 (0.316)**	0.847 (0.335)**
Tournament (Task 2) Performance	0.098 (0.039)***	0.003 (0.045)	0.008 (0.045)
Tournament (Task2)-Piece Rate (Task1)	0.017 (0.048)	0.012 (0.049)	0.023 (0.050)
Guessed Tournament Rank		-0.517 (0.138)***	-0.535 (0.131)***
Eckel-Grossman Gamble Choice			0.138 (0.061)**
Self-Reported Risk (NLSY)			0.191 (0.097)**
N	244	244	244

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is task 3 choice of compensation scheme (1-tournament 0-piece rate)

Instead, our results support our hypotheses, which are based on the belief that Hispanic and black culture differ in ways that alter norms regarding competitiveness. The self-reported measures of confidence and self-efficacy support the experimental evidence: Although women in both samples tend to be less confident, the gap is almost twice as large for the Hispanic sample (two-sided t-test, $p<0.01$ for Hispanics, $p=0.06$ for blacks). Similarly, black women are significantly more likely to believe they are in control of their outcomes than their Hispanic counterparts (two-sided t-test, $p<0.001$), and the gender difference in self-efficacy is only significant for the Hispanic sample (two-sided t-test, $p<0.05$).

Looking at the items focused on cultural variation, we find a notable difference in family structure. Our black participants are close to three times as likely to come from a single mother home (Fisher's exact, $p<0.001$) and fifty percent more likely to report that a female led their household (Fisher's exact, $p<0.01$). Similarly, our black participants, particularly females, tend to have maternal figures with significantly higher education (two-sided t-test, $p<0.001$) while no such

difference exists for paternal education (two-sided t-test, p=0.89). Taken together, these survey measures indicate that relative to the Hispanic sample, our black participants come from a culture of strong female role models.

To examine how these differences effect competitive preferences, we repeat our probit analysis including participant responses to the survey as well as their interactions with ethnicity and gender (table 2.10)[5]. Both the effect of being female and being a black female are no longer significant implying that our cultural measures capture at least some of the differences between our groups that produce differences in competitive preferences. Black women with single mothers or female heads of households are more likely to compete while the opposite is true for Hispanic women. This is consistent with the hypothesis that the impact of these factors depends on the cultural view of them.

Table 2.10: Repeated Probit Analysis of Task 3 Competition Entry with Beliefs about Relative Performance, Measures of Risk Aversion, and Cultural Controls

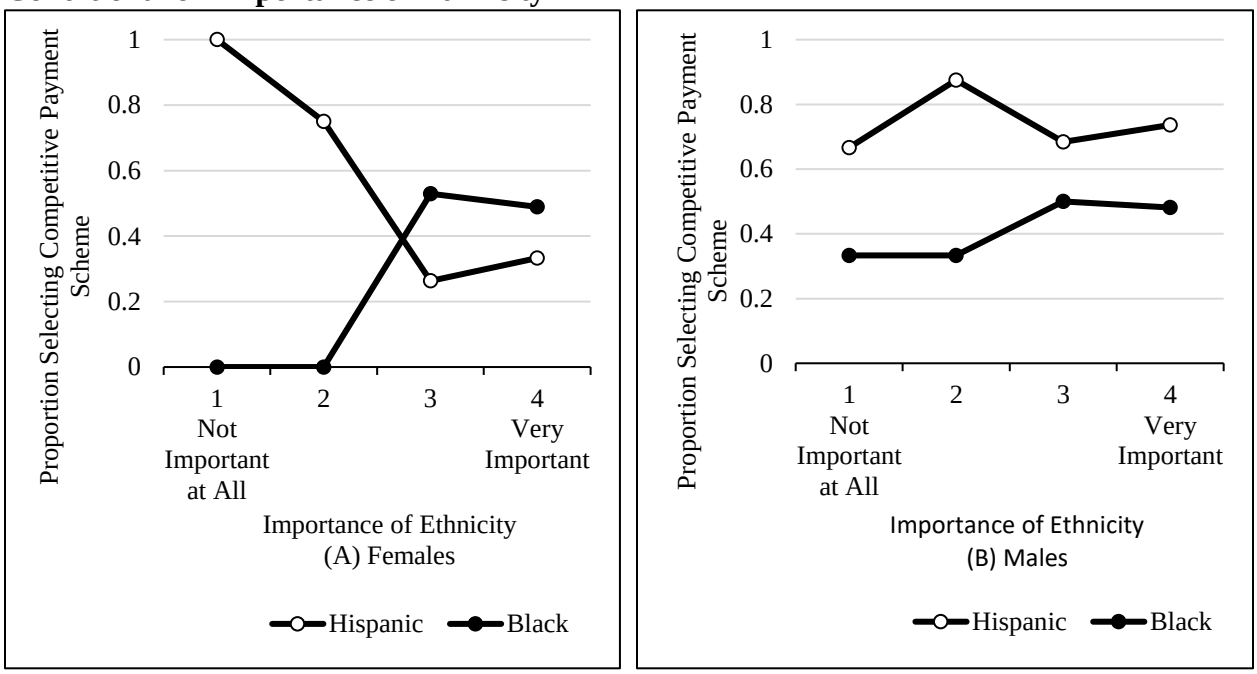
	(1)	(2)
Female	-0.631 (0.283)**	-0.449 (0.388)
Black	-0.540 (0.228)**	-0.544 (0.241)**
Black x Female	0.847 (0.335)**	0.280 (0.454)
Tournament (Task 2) Performance	0.008 (0.045)	0.016 (0.046)
Tournament (Task2)-Piece Rate (Task1)	0.023 (0.050)	0.021 (0.049)
Guessed Tournament Rank	-0.535 (0.131)***	-0.584 (0.134)***
Eckel-Grossman Gamble Choice	0.138 (0.061)**	0.130 (0.065)**
Self-Reported Risk (NLSY)	0.191 (0.097)**	0.130 (0.065)**
N	244	244
Cultural Controls	No	Yes

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is task 3 choice of compensation scheme (1-tournament 0-piece rate)

[5] A Version of the table which includes all of the cultural controls and their interactions can be found in the appendix.

If culture explains our pattern of results, then it is possible that we detect such differences within our samples as well. Our model predicts that Hispanic women who identify more strongly with their ethnicity would be more likely to follow the associated gender norm to avoid competition while the opposite would hold true for black women. In a final exercise looking directly at group membership, figure 2.9 presents the proportion of subjects entering competition based on how strongly they identify with their ethnicity. Participants responded to a prompt asking how important their ethnicity was to them with responses ranging from "1: Not at all" to 4:"Very Important." The black sample identified strongly with their ethnicity: 73% of the black sample said that their ethnicity was very important to them while the same occurred only 45% of the time in the Hispanic Sample.

Figure 2.9: Proportion of Participants Selecting Competitive Payment Scheme in Task 3 Conditional on Importance of Ethnicity



Looking at the female (left) panel of figure 2.9, the pattern of entry supports our hypotheses regarding culture: the more connected a Hispanic female feels to her ethnic identity, the less likely she is to enter competition while the opposite is true for black females. In contrast but consistent with the hypothesis that cultural norms matter, the male panel shows no discernible difference in entry based on ethnic affiliation. Since we hypothesize that males are encouraged to compete independent of culture, it follows that strength of ethnicity should not have a large impact on competitive preferences. In line with the theory, we find the stronger an individual associates with their group, the stronger the inclination to follow that groups behavioral prescriptions. Although we cannot control the treatment of culture as we would other treatments in the laboratory, our survey provides compelling evidence that these cultures differ and these differences affect the existence of a gender gap in competitive preferences.

2.6 Conclusion

Although researchers have identified and replicated the presence of a gender gap in competitive preferences over the past ten years, it is an open question what drives the observed gap. Some authors have found evidence in favor of the role biology may play in the development of these preferences, but others have pointed to the importance of nurture. We add to this literature by comparing two populations that differ solely in culture and find a stark contrast in competitive preferences. We show that our Hispanic sample exhibits the standard competition gap, but we find no such gap exists among our black sample. Further, we illustrate that alternative explanations such as selection or confidence cannot explain our pattern of results. In addition, our survey measures highlight the existence of cultural variation in our samples and that this cultural variation contributes the formation of competitive preferences. Taken together, this paper illustrates the importance of nurture in the development of competitive preferences.

In a broader sense, our work joins others such as Benjamin et al. (2010) in finding that preferences are not determined solely at the individual level but instead that our identity and the groups to which we belong play a role in the formation and expression of preferences. Similar to their work, our results highlight how identity such as gender or culture can systematically influence preferences. However, our work also joins others such as Gelfand et al. (2017) who find that the effect of norms on behavior is often context and environment dependent. This paper contributes to research on identity by highlighting the importance of multiple identities in understanding the strength of specific group norms. Our work highlights how one identity may moderate the effect of another identity's group specific norms on behavior. We find that the influence of identity, specifically gender, on behavior is not the same for all members but instead depends on its interaction with other relevant identities such as ethnicity. Our work highlights the importance and value of considering multiple identities in deepening our understanding of the strength and importance of norms in forming preferences and their resulting behavior.

Policymakers interested in reducing the gender gap in labor market outcomes should focus on campaigns targeting social norms. Although equality is closer to reality than it was a century ago, it is difficult to argue that different social groups receive equal treatment by society. Our work adds to a growing literature that shows identity affects behavior through sending signals about what individuals ought to do. As long as the signal remains that women are not meant to compete, we believe it unlikely if not impossible to achieve gender equality in labor market outcomes. Although shifting social norms is daunting, promising campaigns such as, #likeagirl, have already begun work on shifting the view that female competitiveness is a negative trait. We hope that our work encourages other organizations to pursue similar campaigns that encourage norms focused on the idea that identity does not determine who you are.

3. TRUST AND BETRAYAL: AN INVESTIGATION INTO THE INFLUENCE OF IDENTITY

3.1 Introduction

Although evolutionary biologists have suggested that the formation of social groups may have been essential to survival, group identity has been a source of constant conflict throughout modern human history. The innate tendency to categorize individuals into groups has naturally led to discrimination and in extreme cases the generation of groups such as the Islamic State in Iraq and Syria (ISIS), which explicitly preach the destruction of the out-group. Many major and minor conflicts have had their roots in identity including World War II, The Rwandan Genocide, and more recently, the shooting at the Tree of Life synagogue in Pittsburgh; yet, this list does not begin to scratch the surface of the number of deaths rooted in identity conflict. These instances may differ in the specific reason for conflict, but in most, if not all cases, misunderstanding and mistrust play a key role.

In a world in which the internet has allowed globalization to thrive, companies and individuals face a growing number of interactions with customers and employees outside of their social circles. Although in some cases this has led to greater understanding and empathy among social groups, it has also forced cultures to interact precipitating conflict among contrasting viewpoints that intensifies group divides. Beyond this, recent years have seen the rise of global problems such as global warming and nuclear proliferation requiring cooperation among diverse groups to establish global solutions. Ideally, individuals' identities would not factor into the decision to interact with others; however, a growing body of literature has found that this is not the case (Fershtman and Gneezy, 2001; Fershtman et al. 2005; Eckel and Wilson, 2006). Given

the costs of identity conflict and the benefits of cooperation, it is necessary to understand what drives observed in-group bias.

Akerloff and Kranton's (2000) seminal work on identity economics provides a framework for understanding how identity may influence actions and discusses how such a framework can be used to explain a series of befuddling findings. Their work highlights the importance of group identity in explaining behavior others would deem irrational including costly conflict. In line with this theoretical work, researchers have shown that given the option, participants prefer to enter trusting relationships with members of their in-group even if it is costly to do so (Fershtman, et al. 2005; Falk and Zehnder, 2013; Banuri et al. 2012). This unfair treatment of the out-group and nepotism towards the in-group is observed as taste-based discrimination which causes disparity in opportunity and economic outcomes across social groups. We seek to add to our understanding of this preference by investigating the mechanism behind in-group bias in trust.

To do so, we expand the experimental design first introduced by Bohnet and Zeckhauser (2004) in which the authors examine how the decision to trust differs from a similarly risky decision. The authors identify two major departures in that there is another individual who may earn more than the decision maker and the outcome of trust is determined by another individual as opposed to chance. The researchers found that the latter causes individuals to be significantly less likely to take a risk, and coin this motivation, betrayal aversion. We build on this work by investigating whether the cost of betrayal varies with the identity of the betrayer and how this may explain the presence of in-group bias in trust.

On the surface, we replicate others in finding the presence of in-group bias in trust. On average, participants require the chance of their trust being reciprocated to be roughly ten percent higher for out-group rather than in-group members before entering a trusting relationship. We find

that this premium is driven by betrayal concerns rather than concerns over relative earnings. Our results provide deeper understanding of in-group bias in trust and how it may be reduced to encourage intergroup cooperation. Our work highlights that the observed in-group bias in trust is due to emotional considerations related to betrayal and the desire to conform to group ideals rather than concerns over potential disadvantageous inequality in earnings. With this in mind, policy makers should focus on policies that encourage intergroup cooperation through shielding individuals from potential betrayals or decreasing the cost in terms of group affiliation of trusting an individual from another group.

3.2 Literature Review

Tajfel and Turner (1979) provide the first theoretical work on the effects of group membership. Synthesizing a number of previous experimental results, they argue that the mere recognition of two distinct groups (social categorization) is sufficient to produce intergroup discrimination. They define this tendency to favor members of one's group as in-group bias. Turner et al. (1979) introduce the concept of the minimal group to study the general effect of group membership. In their study, participants were lead to believe they are assigned to groups based on painting preferences. The authors found that this categorization alone was sufficient to cause participants to feel attached and to hold positive beliefs about their group. Beyond this, the authors found that participants tended to give more money to members of their group when dividing a set amount of money among other participants.

Since their original publication, a number of experimental economics studies have attempted to incorporate the minimal group paradigm to study the influence of identity on individual behavior. Eckel and Grossman (2005) used a number of group assignment methods to test when group identity matters and how it effects behavior. They found that categorization itself

is not enough, but when group members interact, they later donate significantly more to a public good that benefits the in-group. Chen and Li (2009) expanded on this work by testing a similar range of group assignments using a menu of games first employed in Charness and Rabin (2002) to identify the effect of group identity on distributional preferences. They found that individuals are more likely to pick social welfare maximizing actions, are less like to exhibit envy and are less likely to punish when their partner is a member of the same group. Building on this work, Chen and Chen (2011) and Li and Liu (2017) find evidence that more efficient equilibria can be reached when in-group matching occurs with artificial group assignment. Using real groups, Holm and Nystedt (2005) and Falk and Zehnder (2013) find that individuals tend to exhibit greater trust in members of their in-group and that this trust leads to more socially desirable outcomes. Taken together, this body of research illustrates that identity can be effective in promoting social welfare, but only in the case in which individuals are interacting with other members of their group.

Although identity can be effective in encouraging socially desirable behavior with in-group members, it can lead to inefficient outcomes in intergroup interactions. One such example is the work of Berhnhard et al. (2006) in which indigenous tribes in Papua New Guinea play a classic dictator game with punishment. The authors found that participants are more lenient towards in-group norm violators and apply more punishment when the victim of the violation is an in-group member. Goette et al. (2006) exploit random assignment of soldiers to platoons in the Swiss army to study the effect of group identity on cooperation and punishment, and found that their participants were more likely to cooperate with members of their platoon and less likely to punish defection from fellow platoon members. Both experiments find that individuals are less likely to punish members of their in-group, which may imply the emotional cost of in-group betrayal is lower (or at least that harsh punishment would lead to greater identity based utility losses) as it

may influence the amount of punishment that is levied. Looking directly at trust, Banuri et al. (2012) find that participants prefer to enter trusting relationships with members of their in-group even if it is costly to do so. Taken together, this research illustrates that identity can encourage cooperation among the in-group but can also lead to unfair and inefficient treatment of out-group members. Our work expands on this literature by investigating the mechanism behind the previously found in-group bias in trust.

To do so, we expand the design first introduced by Bohnet and Zeckhauser (2004) in which the authors investigate how the decision to trust differs from a similarly risky decision. The authors identify the existence of betrayal aversion, or that participants are significantly less likely to take a gamble when another person as opposed to chance determines its outcome. In follow-up work, Bohnet et al. (2008, 2010) illustrate that the presence of betrayal aversion is robust as it appears in a large number of countries that vary widely in culture. Introducing an individual level betrayal aversion measure, Aimone and Houser (2012) show that the potential for betrayal leads to a 35% drop in the number of individuals willing to take a risk, but in later work, that betrayal aversion can be beneficial in increasing trustworthy behavior (Aimone and Houser, 2011, 2013). Additionally, Aimone et al. (2014) show that the decision to trust which includes a potential betrayal activates different areas of the brain (generally associated with emotion) than a similarly risky decision with no betrayal potential. Taken together, this body of research illustrates that the decision to trust is different from other risky decisions and this difference is driven by the possibility of realized betrayal.

Most similar to our work, Hong and Bohnet (2007) investigate how betrayal aversion may manifest differently in different real social groups. They conclude that the level of betrayal aversion is directly related to social status in that high status groups tend to exhibit more betrayal

aversion. Their work documents the presence of different levels of betrayal aversion across social groups but does not measure potential differences in betrayal aversion caused by interactions between social groups. Beyond this, their work illustrates that general in-group bias may be difficult to identify because of differences across real groups in status and norms. To avoid this potential confound, we exploit random assignment to dormitories at Rice University and elicit participants willingness to risk a betrayal conditional on whether the potential betrayer is a member of their in-group. In doing so, we can cleanly identify how the shared identity of a potential betrayer affects the propensity to trust and in a larger sense how betrayal aversion affects intergroup interaction. Our work adds to our understanding of betrayal aversion and trusting behavior in general by being the first to cleanly isolate the mechanism underlying general in-group bias in trust.

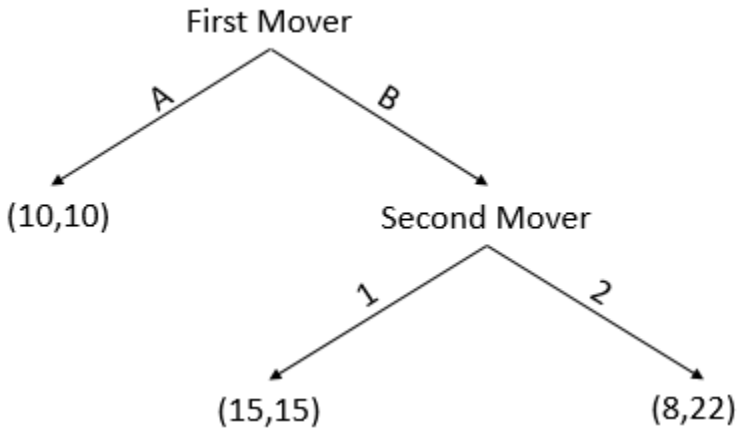
3.3 Experimental Design

3.3.1 The Original Trust Game

The treatments themselves are all variations on the aptly named "Trust Game" introduced by Berg, Dickhaut, Mccabe (1995). In the original experiment, a first mover must decide how much of a $10 endowment to send to a second mover in a one-shot interaction. Any amount sent to the second mover is tripled and then the second mover has the option of returning any of the tripled amount to the first mover. Standard game theory with payoff maximizing agents predicts any rational second mover will not return any money and, anticipating this, any first mover will not send any money, leading to the Nash prediction of zero sent and zero returned. Yet, the original experiment, and the body of experimental literature that it spawned, have shown that first movers trust, regularly sending an amount greater than zero, and second movers reciprocate, tending to return amounts greater than the amount sent.

Following Bohnet and Zeckhauser (2004), we adopt a binary version of this game in which first movers decide whether to trust a second mover, while the second movers decide whether they want to reciprocate in the case that they are trusted. The general structure of the game is depicted below in figure 3.1.

Figure 3.1: The Dyadic Trust Game



As illustrated by the figure, first movers choose "A" or "B." Choosing "B" is similar to sending the entire endowment of $10, or trusting the second mover, whereas choosing "A" is similar to sending nothing. We elicit second mover's preferences using the strategy method: second movers choose "1" or "2" to be implemented in the case that their first mover picks option "B." Choosing "1" is akin to returning more than what was sent, reciprocating on the first mover's trust, whereas choosing "2" is similar to the second mover choosing to send back less than the amount that was sent to them.

3.3.2 Minimal Acceptable Probabilities (MAPs)

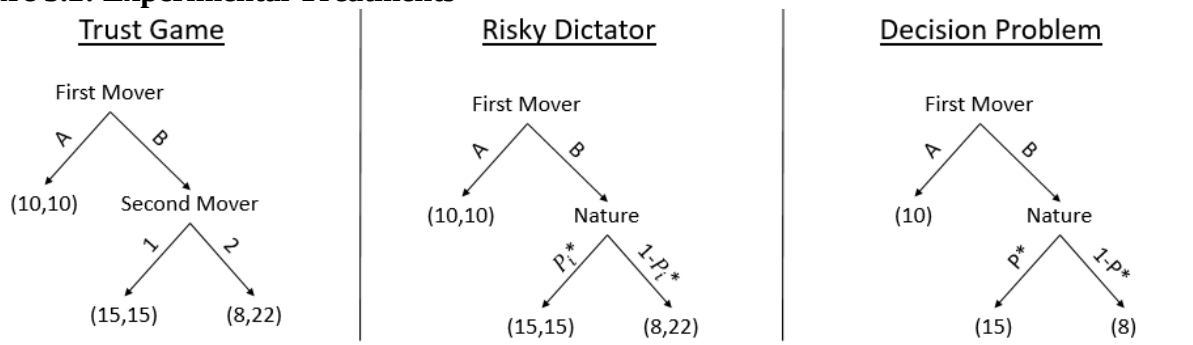
Following Bohnet and Zeckhauser (2004), we adopt an elicitation method similar to the

strategy method for first movers. First movers face a similar decision to that depicted in figure 3.1. However, they make this decision conditional on the probability of receiving the higher payoff in B: first movers provide their minimal acceptable probability (MAP), the lowest chance of receiving the higher payoff at which they prefer option B over option A. Participants are informed that we implement option B whenever their MAP is lower than the actual probability of receiving the higher payoff, i.e. if a participant picks a MAP of 50% then we will only implement option B if the actual probability of receiving the higher payoff is greater than or equal to 50%, otherwise we implement option A. Adopting this approach allows us to avoid concerns about the direct role beliefs[6] may play in behavior as the actions of a first mover in the standard trust game are likely to be related to their subjective beliefs about the trustworthiness of their counterpart (Goette et al. 2006; Fershtman and Gneezy, 2001; Eckel and Petrie, 2011). We avoid this concern by allowing participants to express their preferences for every possible level of trustworthiness. Additionally, this elicitation (combined with the separate treatments) allows for a finer identification of the specific mechanisms underlying trust by allowing more variation in participant's elicited degree of trust.

Figure 3.2 depicts the variations of the binary trust game used as treatments within our experiment. The first treatment, the trust game, is our treatment of interest while the other two serve as controls allowing us to cleanly separate various motivations affecting the propensity to trust. Below, we describe the difference between the treatments and their purpose.

[6] Aimone and Houser (2012) discuss a venue through which beliefs may effect MAPS. We discuss this possibility in the discussion section.

Figure 3.2: Experimental Treatments



P_i * represents the probability that Nature selects the higher first mover payoff when paired with a second mover from group *i*
Note: In the actual conditions, first movers choices are dependent on the probability that they receive option 1. Subjects decide their minimal acceptable probability (MAP) of receiving option 1 at which they prefer option B over option A and B is implemented whenever the actual chance (P*) of receiving the higher first mover payoff is greater than the participant's MAP.

3.3.3 The Trust Game Treatment

At the start of each session, we randomly assign participants to be either the first or the second mover. First movers decide the lowest probability (MAP) of being paired with a second mover who picked "1" at which they prefer "B" over "A". Participants are informed that we will implement an option based on their MAP: if the proportion of second movers who selected "1" is greater than their MAP then we implement "B" and "A" otherwise. If "B" is implemented, the outcome is determined by the decision of a paired second mover.

Following Bohnet and Zeckhauser (2004), we intend not to give the impression to second movers that their decisions could affect the probability of their decision being implemented. Second movers are only informed that they are selecting an outcome in the case that their partner picks "B". After making their decision, we inform second movers about the method in which we implement decisions.

We interpret participant's statement of MAP is as the participant's willingness to enter a trusting relationship. We conduct two additional treatments to understand how betrayal aversion contributes to this propensity.

3.3.4 The Risky Dictator Treatment

Similar to the trust game treatment, we randomly assign participants to one of two roles at the onset of the session. The difference between this treatment and the trust game is the way in which the outcome in option B is determined. As before, first movers are asked to state the lowest probability (MAP) of receiving "1" at which they prefer taking "B" over "A". However, in this treatment, nature as opposed to another participant determines the outcome in option B. Participants not in the role of first mover act as passive receivers of the outcome determined by nature and their paired counterpart. Following Bohnet and Zeckhauser (2004), the probability of nature selecting the higher first mover payoff is set equal to the proportion of second movers who selected the higher first mover payoff in the first treatment for comparability.

We compare aggregate MAP in this treatment to that from the trust game to investigate how betrayal aversion affects the propensity to trust. Since the only difference between the two treatments is how the outcome is determined, we interpret larger aggregate MAP in the trust game as participants exhibiting betrayal aversion. We conduct one final treatment to separate the effect of risk aversion and other regarding preferences on the decision to trust.

3.3.5 The Decision Problem Treatment

Similar to the previous treatments, we ask participants to state their desire to take option B in the form of the lowest probability of receiving $15. However, this treatment differs in the lack of a second player; the decisions made by participants in this condition only affect their own payoffs.

We compare aggregate MAP from this treatment to that from the risky dictator condition to determine how other regarding preferences (altruism, inequality aversion, etc.) affect the propensity to enter a trusting relationship. We interpret the difference in aggregate MAP between

the two treatments as participants expressing preferences over the earnings of others and efficiency. We interpret MAP from this treatment as a pure measure of risk preferences as participants are expressing their desire to take a gamble over a sure option. Thus, any MAP greater than the risk neutral point[7] is interpreted as participants exhibiting risk aversion.

Taken together, the three treatments allow us to isolate separate motivations to trust. Aggregate MAP from the trust game provides a measure of the overall propensity to trust. We compare aggregate MAP from the trust and risky dictator games to identify how betrayal aversion contributes to the propensity to trust. We compare aggregate MAP from the risky dictator game with the decision problem to identify how concerns over others' earnings affects the propensity to trust. Lastly, the aggregate MAP from the decision problem allows us to identify how risk aversion contributes to the propensity to trust.

3.3.6 Ex-post Survey

Following each treatment, each participant completed a short survey. We designed the survey to control for various concerns that may affect our results. The survey included a measure of generalized trust, various relevant demographics, and the strength of one's affiliation to their residential college.

3.3.7 Random Assignment to Real Identities

To isolate the mechanism behind in-group favoritism in trust, we expand the original design introduced in Bohnet and Zeckhauser (2004) to include the within-subjects factor of identity. Participants still experience only one of the treatments described above but in the treatments where another player is present (trust and risky dictator), participants make two decisions: one with a member of their in-group and another with a member of their out-group. We

[7] A 29% chance of receiving $15 would make a risk neutral individual indifferent between the two options

inform participants that the order of decisions as well as which decision is implemented for payment is randomly determined.

Given the work of Hong and Bohnet (2007), we paid close attention to the identities we used to avoid potential confounds in status and norm differentials. We conducted our study at Rice University in the fall of 2016 as it provides an ideal test bed for measuring the general effects of group membership. Incoming freshman have no control over the initial residential college assignment minus a few exceptions[8]. Students spend their first weeks on campus being indoctrinated into their college and must spend meals with fellow members of their college for the duration of their stay; Students can be seen sporting clothes that signify their allegiance to their residential college, and students regularly compete in friendly competition against members of other residential colleges. Similar to Goette et al. (2006), we avoid concerns over status, selection, and cultural norms as relevant group characteristics are evenly distributed among the residential colleges through random assignment. This random assignment is ideal for measuring general in-group bias as the individual groups do not differ in any systematic manner but are "real" in the sense that their identities matter outside of the laboratory. This facet of our design allows us to be confident in our identification of preferences over general intergroup interactions rather than group specific prescriptions given random assignment to real identities or potentially artificial effects created by using minimal groups.

To identify the effects of identity on the propensity to trust, we compare aggregate MAPS within treatments. A difference in the aggregate MAP for in-group and out-group matches in the risky dictator treatment is interpreted as participants having preferences over others' earnings

[8] These exceptions include students with disabilities and students whose parents graduated from a specific residential college (Rice University Website, 2018)

which are dependent on the identity of their counterpart. Similarly, a difference in aggregate MAP for in-group and out-group matches in the trust game provides a measure of how identity affects the propensity to trust. We subtract the aggregate MAP for in-group matches from the risky dictator game from that of the trust game to provide a measure of betrayal aversion towards the in-group. We compare this to the same measure for out-group matches to investigate how betrayal costs and their impact on the propensity to trust may differ by counterpart identity.

We ran 9 sessions during the fall of 2016. We recruited participants from each of the 11 residential colleges present at Rice with each session containing students from no more than three colleges. Sessions were run on the top floor of the Library in the Behavioral Research Lab created by Dr. Rick Wilson. The experiment was run on computers using ZTREE (Fishbacher, 2007) and each session lasted approximately 45 minutes. Participants earned an average of $16 including a $5 show-up fee. We had 26 pairs in the first treatment across 4 sessions, 27 in the second treatment across 4 sessions, and 16 participants in the third treatment.

3.4 Theory and Hypotheses

Our hypotheses about aggregate (non-identity based) behavior in our treatments are based on the predictions and previous results provided by Bohnet and Zeckhauser (2004).

Hypothesis 1: Individuals will behave altruistically observed as a lower aggregate MAP in the risky dictator treatment relative to the decision problem treatment.

The support for this hypothesis comes from a number of previous works on other regarding preferences. Perhaps the strongest support comes from the original work in which the author's identify that individuals report a lower aggregate MAP in the risky dictator relative to the decision problem treatment (Bohnet and Zeckhauser, 2004). More general support comes from Charness and Rabin (2002) where the authors run a series of games to identify other regarding preferences

and find that individuals tend to be more concerned with social welfare than reducing differences in relative earnings.

Hypothesis 2: Individuals will be betrayal averse measured as a higher aggregate MAP in the trust game treatment relative to the risky dictator treatment.

The support for this hypothesis comes from the original paper as well as the large body of research it has inspired. Papers in this framework have consistently identified that presence of betrayal aversion as measured as a difference in aggregate MAP between risky dictator and trust game treatments (Bohnet and Zeckhauser, 2004; Hong and Bohnet, 2007; Aimone and Houser, 2012).

Taken together, these hypotheses simply state that we expect to replicate the original results. However, the purpose of this paper is to expand the original design to learn about how these preferences may be affected by identity. We expand our predictions to include identity using the theoretical framework outlined in Akerloff and Kranton (2000).

We focus on how identity may prescribe what one ought to do in interpersonal interaction and how the cost of betrayal may vary by the group identity of others. This general model provides intuition in two important ways that will inform our hypotheses:

First, One's behavior can result in identity based utility gains and losses. Each identity comes with a set of social prescriptions and actions contrasting such prescriptions are likely to incur utility losses. Ignoring one's group prescriptions can create feelings of disconnect from the group or shame for failing to be an ideal member of the group.

Second, Other's behavior can result in identity based utility gains and losses. Members of one's group failing to conform to identity prescriptions or exhibiting behavior that excludes one from the group can incur a utility loss caused by a loss of identity. Exclusion can create feelings of

disconnect from the group and others failing to hold to a group ideal may decrease the overall value of membership to the group leading to identity based utility losses.

Of note, the intuition above presumes that membership to the group is a positive. If this is not the case, individuals may prefer actions that distance them from the group. We believe that this case is unlikely for our subjects for two reasons: 1. survey responses about their attachment to their residential colleges[9], and 2. unlike membership to some other groups, individuals have the option to leave. Using this framework, we generate identity specific hypotheses:

Hypothesis 3: P*, the probability of being paired with a counterpart who selected the equitable payoff, will be higher for in-group matches than out-group matches. We predict the proportion of individuals reciprocating trust will be higher among in-group decisions.

We assume that membership to a group contains the prescription that one should reciprocate trust if the other party is a member of one's group. Thus, reciprocating trust strengthens membership to the group leading to identity based utility gains while failing to do so would incur identity based utility losses. As a result, not reciprocating on the trust of one's in-group incurs an additional cost that is not present in a similar decision with an out-group member. Since the costs (benefits) of betrayal are higher (lower) among one's in-group, we expect participants to be less likely to betray when the betrayal affects a member of their in-group.

Hypothesis 4a: Individuals will act altruistically towards in-group matches. Aggregate MAP for in-group matches in the risky dictator treatment will be lower than aggregate MAP from the decision problem treatment.

[9] Mean response to a survey item rating "I am happy with my residential college" was 4.32 out of 5 where 5 is "strongly agree"

A choice that results in higher earnings for the in-group carries additional benefits as it strengthens one's group affiliation while simultaneously increasing the value of that membership through higher earnings for the group. Given utility is affected by the groups to which the individual belongs to, a choice that improves welfare for the group also increases welfare for the individual.

Hypothesis 4b: Individuals will act inequality averse towards out-group matches. Aggregate MAP for out-group matches in the risky dictator treatment will be higher than aggregate MAP from the decision problem treatment.

A choice that results in higher earnings for an out-group member weakens the decision maker's group affiliation while simultaneously weakening the value of that group membership as an out-group's status is raised. Given utility is affected by the groups to which the individual belongs to, a choice that improves welfare for the out-group makes the in-group relatively worse off which carries to the decision maker as well.

Hypothesis 4c: Individuals in the risky dictator treatment will be more likely to prefer the gamble when they are matched with an in-group member. Aggregate MAP in the risky dictator treatment will be larger for out-group matches.

This hypothesis can be viewed as the direct result of the two preceding hypotheses. Our model predicts that individuals have additional incentives to take the gamble with the in-group and to take the sure payment with the out-group. Picking a lower MAP for an in-group match affirms one's membership to the group and increases the in-groups overall welfare producing identity based utility gains while the opposite is true for out-group matches. Taken together, we expect the aggregate MAP for in-group matches in the risky dictator treatment to be lower than aggregate MAP for out-group matches.

Hypothesis 5a: Individuals will act more betrayal averse towards their in-group as opposed to their out-group.

Being betrayed incurs an emotional cost and this emotional cost is higher when the betrayal comes from a member of the in-group because of the loss of identity. Put in real terms, being betrayed by someone who is close to you carries additional hurt: not only do you experience the betrayal itself, but additional costs arise from how it affects the relationship you share with the individual. You lose trust in the individual, and the group to which you both belong. Past interactions (memories) that were viewed positively may be masked by the betrayal and prevent the formation of future positive interactions. In our model, this is precisely the loss of identity predicted by individuals picking actions that exclude one of its group members.

Hypothesis 5b: We expect individuals to act more betrayal averse towards their out-group as opposed to their in-group

Individuals are more likely to pick actions that benefit the in-group because they strengthen their own group affiliation. In this way, individuals may be motivated to report lower maps for the in-group because it affirms their membership to the group. One possibility is that this motivation is already captured by the risky dictator treatment, but this decision requires trust. Choosing to trust is a stronger action than those in the risky dictator treatment and therefore may exacerbate the observed difference in decisions between in-group and out-group matches.

An alternative explanation for this hypothesis is that what constitutes a betrayal may be different for in-group and out group members. Previous experimental evidence involving in-group and out-group interactions provide tangential support for this claim. Researchers have shown that individuals are less likely to punish violated expectations when the violation comes from an in-group member (Goette et al., 2006; Bernhard et al. 2006) and more recently, that individuals are

more likely to return to cooperation after a defection when they repeatedly interact with in-group members as opposed to out-group members (Li and Liu, 2017). In both of these examples, participants choose to avoid actions that would further reduce their group identity even though such behavior may be warranted and readily apparent in other circumstances.

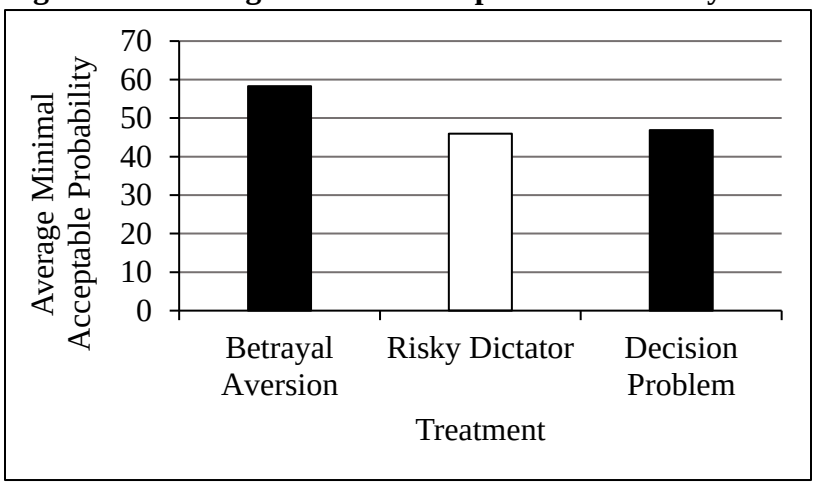
Taken together, we expect individuals to report lower aggregate MAPs towards in-group matches in the risky dictator treatment while bias could occur in either direction towards the in-group in the trust game treatment. In line with previous work on trust and identity, we expect aggregate MAPs to be lower for in-group matches in both the risky dictator and trust game. Yet, it may be possible that betrayal aversion is lower towards the out-group even in this case as betrayal aversion requires evaluating MAPs relative to aggregate behavior in the risky dictator game.

3.5 Results

3.5.1 Aggregate Results

Before looking at identity based MAPs, we compare our aggregate results to previous work. Figure 3.3 presents aggregate Map across all treatments and in the treatments where individuals make decisions with in-group and out-group matches, MAPs are pooled.

Figure 3.3: Average Minimal Acceptable Probability across Treatment



To measure general risk aversion, we compare the aggregate MAP in the decision problem with the risk neutral point ($p(15) = 0.29$). On average, participants in our sample require a 46.88% chance of getting 15 tokens from the gamble before opting for the gamble over the sure option; this difference is statistically different from the risk neutral point of 29% (two-sided t-test, $p<0.01$). Compared to the original results (Bohnet and Zeckhauser, 2004), our participants are more risk averse (46.88% vs. 37%) but not significantly so.

We compare the aggregate map in the risky dictator treatment with that in the decision problem to measure inequality aversion. Looking at figure 3.3, participants in the risky dictator treatment required a 46% chance of getting the payoff that provides 15 tokens before preferring the gamble. When we compare this to behavior in the decision problem we can see that on average participants are more likely to take the gamble when we introduced the second player (0.88%) but the difference is not statistically different from zero (two-sided t-test, $p=0.78$). We interpret this as participants having low concerns over relative earnings and efficiency in this environment. Bohnet and Zeckhauser (2004) similarly detect no difference in behavior between the decision problem and risky dictator game but their participants have an average MAP about 10% lower than our participants.

To test for general betrayal aversion, we compare aggregate behavior in the trust game with that in the risky dictator treatment. On average, participants require a 58.29% chance of getting 15 tokens before preferring option B when another individual determines the outcome. Compared to the risky dictator treatment, participants are requiring a 12.3% higher chance when there is a chance that they are betrayed by another human; this difference is statistically significant (two-sided t-test, $p<0.01$). The degree of betrayal aversion is smaller in magnitude than the original

paper but we reach the same qualitative conclusion that individuals are averse to learning they have been betrayed.

Overall, we confirm our first two hypotheses and replicate the pattern of findings found in Bohnet and Zeckhauser (2004). However, the main purpose of this paper is to understand how identity influences the decision to trust. Thus, we compare the decisions towards in-group and out-group members in the risky dictator and trust game treatments. First, we compare the rate at which second movers reciprocate on trust towards first and second movers.

3.5.2 Identity Results

Figure 3.4: Second Mover Behavior by Counterpart Identity

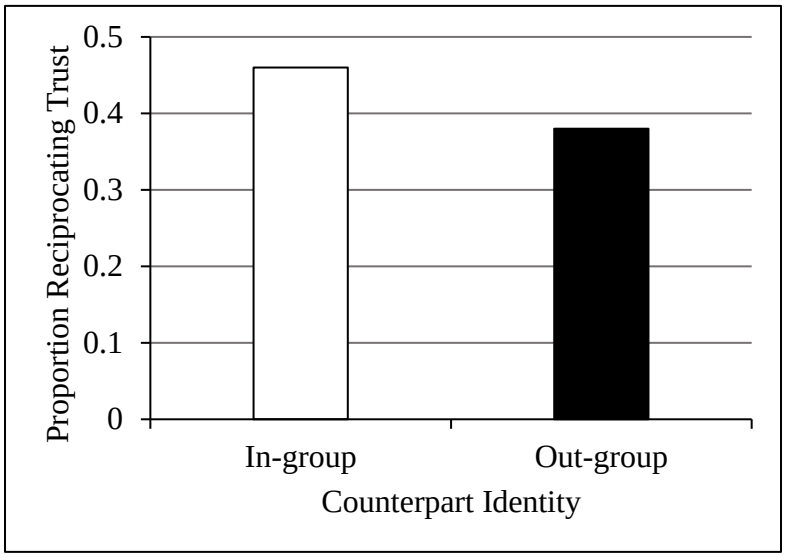
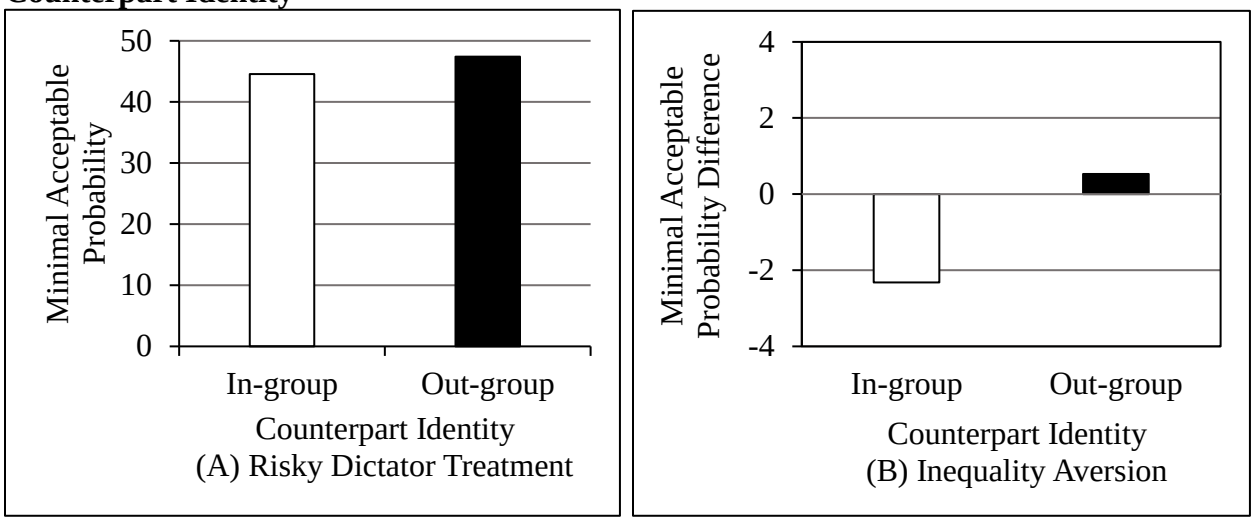


Figure 3.4 presents the proportion of second movers choosing to reciprocate on the trust of their first mover by counterpart identity. Second movers reciprocate 61.5% of the time when matched to an out-group member and 53.8% when matched with an in-group member. We find that second movers are more likely to reciprocate on trust when paired with a member of their

residential college but the difference is not significant (two-sided test of proportions, p=0.57). Directionally, this is consistent with our hypothesis albeit insignificantly so.

Figure 3.5: Average Minimal Acceptable Probability in the Risky Dictator Treatment by Counterpart Identity



Turning to the behavior of first movers, Figure 3.5 presents the aggregate in-group and out-group MAP in the risky dictator treatment. The panel on the left presents the aggregate MAP while the panel on the right presents the results accounting for risk aversion. The left panel indicates our participants tend to show slight favoritism towards their in-group requiring a 2.85% higher chance of getting 15 tokens when they are matched with a member of another residential college. When we control for the aggregate level of risk aversion we find that individuals are actually more likely to take the gamble when an in-group member is introduced while the opposite holds true for out-group matches. Directionally, this is in line with our hypotheses but the differences are small and not significantly different from zero in either case (two-sided t-test, p=0.32).

Figure 3.6: Average Minimal Acceptable Probability in the Betrayal Aversion Treatment by Counterpart Identity

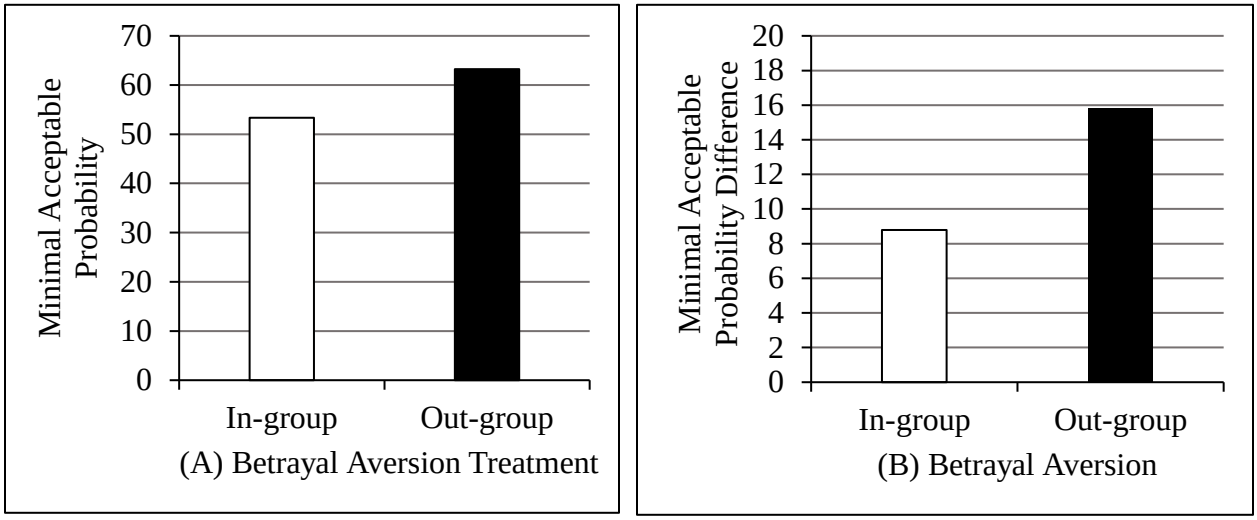


Figure 3.6 presents the MAPs from the trust game by counterpart identity. The left panel displays the aggregate MAPs while the right presents the aggregate MAPs controlling for risk aversion and inequality aversion. Looking at the overall in-group bias in trust (left panel), we find that on average individuals require about a 10% higher chance of having trust reciprocated before entering a trusting relationship with an out-group member as opposed to an in-group number (two-sided t-test, p=0.06). The right panel presents the aggregate MAPs after controlling for risk aversion and other regarding preferences. We find that participants are significantly betrayal averse towards in-group and out-group members. This result adds to the extensive literature finding the presence of emotional costs associated with betrayal. When we compare the level of betrayal aversion between in-group and out-group matches, we find that the implied cost of betrayal is almost twice as high for out-group matches as opposed to in-group matches.

Table 3.1: Regression Estimates of Minimal Acceptable Probability

	(1)	(2)	(3)
Trust Game	17.62 (5.18)***	15.32 (8.76)*	16.95 (5.37)***
Trust Game x In-group	-9.88 (3.59)***	-9.88 (3.71)***	-9.88 (3.79)**
Risky Dictator x In-group	-2.85 (2.85)	-2.85(2.94)	-2.85 (3.00)
Male	-9.17 (5.17)*	-8.45* (4.97)	-8.19 (5.18)
In-Group First	-5.91 (5.54)	-4.04 (5.89)	-5.66 (5.99)
Cons	61.39(8.24)***	61.56 (9.13)***	61.13 (8.65)***
Num of Obs	106	106	106
N	53	53	53
Session Fixed Effects	No	Yes	No
College Fixed Effects	No	No	Yes

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the subject level
Dependent variable is minimal acceptable probability
Note: Decision problem excluded to include in-group dummies

To analyze our data in further detail, we adopt an Ordinary Least Squares (OLS) regression with individual MAPS as the dependent variable (table 3.1). In each specification, we drop observations from the decision problem allowing us to investigate the interaction of in-group pairing on MAP decisions. Across all three specifications, we can see there is a significant effect of having another individual decide the outcome of the gamble. We interpret the coefficient by comparing to the omitted category, the risky dictator treatment. Thus, on average participants are requiring a greater than 15% increase in the chance of getting the higher payoff before taking the gamble when it includes the potential for betrayal. We also find evidence that the difference in betrayal costs associated with identity are robust and statistically different from zero. Participants require on average a close to 10% increase in the probability of getting the good outcome before risking betrayal from an out-group member as opposed to an in-group member. We find support for our hypotheses that individuals will exhibit higher betrayal aversion towards the out-group.

3.6 Conclusion

Our results provide mixed confirmation of our hypotheses. We find significant evidence that individuals prefer to enter trusting relationships with members of their own group and that individuals reciprocate trust more when paired with an in-group member. However, first movers tend not to display concerns over relative earnings that depend on the identity of their counterpart. Instead, we find that in-group bias in trusting behavior is driven by the presence of the possibility of betrayal.

In our theoretical framework, this result implies that individuals care more about how their own actions effect their group membership rather than the actions of others. One potential explanation for our results is the ability to divorce others actions from their group association: the choice of minimal acceptable probability cannot be divorced from group membership as it reflects one's desire to conform to group prescriptions while another's choice to betray can be as it may reflect on them instead of the group. Put another way, the identity-based cost of not trusting the in-group may be higher than the cost imposed by any realized betrayal.

An alternative explanation is participants hedging against a potential betrayal: any minimal acceptable probability besides 100% entails the possibility of betrayal and providing a lower minimal acceptable probability may be an attempt to pick an action that creates additional identity-based utility gains that can offset any losses produced by a realized betrayal. This would be akin to affirming one's membership to a group by signaling desire to be a member of the group independent of negative treatment by other members.

A third potential explanation is that the cost of betrayal is higher when it comes from the out-group in the scenario we have studied. This would be the case in which the rarity of interactions with the out-group places more weight on potential out-group betrayals or may be potentially

related to the stakes of the trusting decision. When individuals think about betrayal from the in-group they may gravitate towards larger betrayals (such as infidelity or slander) that are related to the degree of trust extended to in-group members whereas the same consideration for the out-group may be more analogous to the situation we studied in the lab. Put in real terms, this would be similar to being willing to overlook a small infraction from an in-group member but avoiding the same potential for betrayal from an out-group member.

Of note, Aimone and Houser (2012) provide an alternative explanation for behavior in this environment that is related to the interaction of beliefs and loss aversion. The authors suggest that participant's choice of MAP could be affected by beliefs through the desire to avoid finding out that beliefs are incorrect. For example, if a participant believes their in-group is 70% trustworthy but is risk neutral and prefers to take the gamble whenever P*=0.29, then they may report a MAP higher than p=0.29 to hedge against the possibility of finding out that less than 70% of their in-group is trustworthy. We believe these concerns could only strengthen our results: beliefs towards the in-group tend to be biased upwards (Falk and Zehnder, 2013) suggesting that participant's MAPs toward the in-group would be inflated upwards. Thus, if anything, our results regarding the costs of betrayal from the in-group would be strengthened by this concern.

Before concluding, we discuss how our results may generalize to other groups. As part of our procedure, individuals reported the strength of their affiliation to the residential college along a number of dimensions. We find no evidence that any of these measures predict discrimination towards the in-group. However, we believe this result may be an artefact of the identities we are using: residential colleges creates variation in group identity but are housed under the shared identity of being Rice students, which may limit differentiation between the in-group and out-group. Similar to the concerns raised by Aimone and Houser (2012), we believe that our results

would only be stronger with other groups. As groups become more socially distant, we predict the difference in MAPs would as well. However, using other groups introduce a host of other concerns including beliefs about trustworthiness, status, and group specific norms. The strength of our design lies in its ability to avoid these concerns but trusting relationships outside the lab depend deeply on these concerns and researchers should take care in considering how these additional concerns may affect the propensity to trust. Although neither Hong and Bohnet (2007) nor Bohnet et al. (2008) includes explicit intergroup interaction, they provide evidence that the strength of various concerns related to the propensity to trust vary by social group and culture implying that specific group norms matter in this context.

This paper is the first to investigate how the identity of one's counterpart effects betrayal aversion. We exploit random real group assignment to dormitories at Rice to avoid concerns about status and selection. We find that participants have little to no preference over other's earnings whether they are members of the in-group or the out-group, but that participants are betrayal averse towards both. Additionally, we find that the level of betrayal aversion is significantly higher towards the out-group.

In a time when tribalism dominates political discourse and divisive lines are often drawn between groups, we believe studying the determinants of intergroup hostility to be of utmost importance. Our results suggest that individuals are motivated by the emotional rather than monetary aspects of the decision to trust. Policy makers interested in reducing intergroup tension should focus on shielding trustors from potential emotional costs rather than highlight the potential monetary gains from trust. Our results suggest that interaction with the out-group is a necessary first step in reducing intergroup hostility and encouraging intergroup cooperation. Policies could

focus on insuring individuals from potential betrayals or encouraging regular intergroup interaction to reduce the weight placed on the actions of individual out-group members.

Future work on this topic should extend our work to real social groups and should focus on separating out the various explanations for our results. Although our work provides a necessary first step by isolating general intergroup behavior, additional research is necessary to understand how our results may vary with specific group attributes including status and norms. Researcher should focus on groups with varied status and antagonism to understand whether different strategies for encouraging trust are necessary for different social groups. Similarly, while we have identified that identity-based differences in the perceived cost of betrayal are an important factor in the decision to trust, we have not identified why this is the case. We discuss how various explanations may be consistent with our results but future work is necessary to understand the importance of volume of interactions, the relative weights of one's versus another's actions in terms of identity-based utility effects, and whether our results are robust to situations that vary the type of betrayal.

4. GENDER, COSTS, AND LOW PROMOTABILITY TASKS

4.1 Introduction

Although the gender gap in wages has shrunk considerably in the last 50 years, researchers have noted significant vertical segregation across genders (Altonji and Blank, 1999; Bertrand and Hallock, 2001). Historical accounts seeking to explain the persistence of gender differences in labor market outcomes focus on differences in skills, different preferences for working, and discrimination (Arrow, 1973; Polachek, 1981; Goldin and Rouse, 2000). These accounts may explain a small part of the gender disparity, but they do not adequately account for its size and robustness (Blau and Khan, 2017).

Behavioral economists have begun to explore alternative explanations for the persistence of gender differences in labor market outcomes. For example, Babcock et al. (2003) find that women are over twice as likely as men to be apprehensive about negotiating for higher wages. As a result, they find that women ask for and receive 30% less than men when negotiating. Vesterlund et al. (2007) investigate an additional explanation in the form of competitive preferences. They find that men are over twice as likely to opt for a competitive payment scheme even though their performance is statistically indistinguishable from that of their female counterparts. Later work by Buser et al. (2012,2014) finds that competitive preferences predict choice of major and as a result occupation. These findings in tangent with the finding that differences in occupational choice are the single largest contributor to the current gender wage gap (Blau and Khan, 2017) imply that behavioral considerations may explain the persistence of the gender differences in labor market outcomes.

More recently, Babcock et al. (2017) investigate a potential alternative explanation by investigating gender differences in the completion of low promotability tasks. The authors define

such tasks as tasks that need to be completed in an office but do not increase the chances of the individual being promoted. Examples of such tasks include making coffee or copies, taking notes during a meeting or, within academia, serving on committees and advising students. Such tasks can be viewed as an individually produced public good that offers benefit to all but less so to its producer as they must pay the cost of production. Providing anecdotal and empirical evidence, Babcock et al. (2017) find that women are more likely to be asked and to accept such tasks in a number of industries including academia, banking and other mid-level jobs (Misra et al., 2012; Mitchell and Hesli, 2013; Benschop and Doorewaard, 1998; De Pater et al., 2010). To understand this pattern of behavior, Babcock and her co-authors (2017) take to the laboratory to study the allocation of tasks that individuals want to be completed but prefer another individual to complete.

The results from the laboratory confirm the impressions provided by previous empirical work. Across 10 rounds, women volunteered to complete greater than one additional low promotability task. Although this may not appear to be a large difference, it is important to note that this translates to a greater than 10% difference in the propensity to complete tasks with low promotability. In a follow-up to the first experiment, the authors investigate if there are any gender differences in how managers allocate requests to complete low promotability tasks and find a reinforcing effect in that over ten rounds, women receive an average of two additional requests to volunteer to complete the low promotability task. Taken together, women are more likely to be asked and to volunteer to complete low promotability tasks. Babcock et al. argue that this difference could explain a significant portion of vertical gender segregation through lower opportunities for promotion.

To build on this result, we explore the mechanism underlying the observed gender gap in low promotability task completion. We introduce heterogeneous costs of low promotability task

completion into the experimental design used in Babcock et al. (2017) to test the importance of gendered beliefs in this context. The intuition behind this addition is that real differences in opportunity costs should create an implication of who ought to volunteer that is stronger than those potentially implied by gender norms. Put another way, we hypothesize that the introduction of explicit opportunity cost differences will results in common recognition of who should complete the low promotability task. Introducing this cost differential will provide our experiment with a comparable structure to the hierarchies present in most office environments, mimicking the standard that workers in a large office setting are paid differentially based on their abilities (Oi and Idson 1999). We believe these wage differences may explicitly signal differing opportunity costs for performing tasks, providing workers with a tacit understanding of whose time is more valuable and should be devoted to more promotable tasks. We expect to find that introducing a similar cost-based formal hierarchy for performing low promotability tasks may reduce the volunteering gap between genders, simultaneously exposing and suppressing a potential implicit gender-based cost structure that may be driving Babcock et al.'s (2017) original results.

We find mixed results with regards to our hypothesis. In our baseline treatment, we fail to replicate the original results. We find a gender gap that is insignificant and close to 80% smaller than the original study suggesting that our sample may be inherently different from the original study. When we introduce heterogeneous costs of volunteering, we find a gap similar in magnitude to the original study but much larger than our baseline treatment. We find that there is substantial differences across cost types with the gender gap being half as a large among low cost individuals relative to their high part counterparts. However, when we ask participants to act as managers who request other participants to complete a low promotability task, we find that the introduction of

explicit cost differences removes the previously observed gender gap in requests received and, as a result, the gender gap in low promotability task completion.

Our results suggest that the empirically observed gender gap in low promotability task completion may be driven, at least in part, by beliefs that differ across gender. The introduction of explicit cost differences in task completion is salient to participants producing a large difference in task completion across cost types. This suggests that the presence of hierarchy in office settings may alleviate gender differences in low promotability task completion across hierarchical levels. However, we find that among high cost individuals the gender gap is even larger than the original study suggesting that the impact of introducing hierarchy into an office setting may not alleviate gender differences within hierarchical levels. On the other hand, we find that the introduction of explicit cost differences in the presence of managers effectively eliminates any gender difference in requests received and, as a result, any gender differences in task completion. This result suggests that reminding manager of the opportunity costs of their employees (which increase with low promotability task completion) could effectively eliminate the gender gap in low promotability task completion.

4.2 Literature Review

Although our study is framed in terms of the completion of low promotability tasks, the experimental game we use was originally introduced as the "Volunteer's dilemma" (VD) by Andreas Diekmann (1985). The name refers to the situation the experimental game represents in which individuals must simultaneously decide whether to sacrifice to produce a good that benefits everyone. Similar to the low promotability task scenarios described above, the volunteer's dilemma is characterized by a case in which everyone prefers for a collective good to be produced but prefers not to be the person to produce it. Tame examples of the volunteer's dilemma include

household chores (Weesie, 1993) and the decision to produce open source software (Johnson, 2002). However, others have pointed out much more extreme versions including bystanders calling the police when a crime is witnessed by multiple individuals (Rosenthal, 2015), and a soldier diving on a live grenade to save their fellow soldiers (Murnighan et al., 1993).

Our work focuses on a particular version of the VD in which individuals face asymmetric costs of volunteering. Diekmann (1984) provides a game theoretic analysis of this case finding the counter-intuitive prediction that the propensity to volunteer is positively related to the cost of volunteering. Put in context, this would imply that in a situation in which a group of bystanders witness a drowning man, individuals who cannot swim would be most likely to try and save the individual (Diekmann, 1984). Instead, Diekmann (1984) suggests a more plausible solution suggested by Schelling's (1980) concept of a prominent solution in which the universal recognition of who ought to volunteer leads higher rates of volunteering among low cost individuals. This prediction may be consistent with the results of Babcock et al. (2017) if gender norms produce a focal point of who should volunteer in the absence of other signals. If this is the case, then the introduction of asymmetric costs of volunteering should alleviate at least part of the gender gap as long as explicit cost differences are more salient than gender norms.

However, other experimental work suggests the possibility of types of individuals that differ in the propensity to volunteer. Bergstrom et al. (2015) finds that between 15% and 36% of individuals are the type that prefer to volunteer as evidenced by a decision to volunteer before other players have an opportunity to do so. Combined with the suggestion by Babcock et al. (2017) that the observed gender gap in volunteering may be the result of gender differences in preferences, our proposed treatment of explicit cost differences may be ineffective at alleviating the previously observed gender gap if women simply have stronger preferences to volunteer.

Yet, other experimental evidence supports the relevance of focal points in the VD context. Przepiorka and Diekmann (2018) show experimental evidence of high rates of coordination in circumstances where there is a "strong" member of each group that faces lower costs of volunteering. Consistent with this result and most similar to our own work, Healy and Pate (2018) find high rates of volunteering among low cost individuals in an asymmetric volunteer's dilemma. Although their work is not explicitly focused on gender, the authors find no evidence of a gender gap in volunteering lending support to the possibility of gender norms serving as a weak focal point (Healy and Pate, 2018). Other suggestive evidence comes from Babcock et al. (2017) in which no gender gap in volunteering appears when experimental sessions only involve a single gender.

We seek to further this body work by introducing asymmetric costs into the treatments used in Babcock et al. (2017). Although our introduction of asymmetric costs is similar to the incomplete information treatment in Healy and Pate's (2018) work, there are important differences. Beyond the use of different parameters, we study the pattern of requesting behavior in the presence of asymmetric costs which, to our knowledge, has not been previously studied. Our work contributes to the literature on volunteering behavior in an asymmetric VD setting, but is the first to investigate how asymmetric costs effects requesting behavior. Beyond this, our work is explicitly focused on understanding the mechanism underlying the gender gap in Babcock et al. (2017) by investigating how the introduction of asymmetric costs may moderate its magnitude.

4.3 Experimental Design

4.3.1 Treatment 1: Symmetric Volunteer's Dilemma

Treatment 1 is a replication of the first treatment from Babcock et al. (2017). After providing informed consent and taking a seat, participants are informed that they will play a game

for ten rounds. Participants are told that in each round they will be randomly assigned to a group of 3 with the condition that they cannot be re-matched with a member of their group from the previous round. In an individual round, participants are endowed with a single token and have 120 seconds to choose whether or not to invest in a group account[10]. As soon as a group member invests, the round ends and each subject receives an additional token. However, the participant who invested must pay a cost of 0.75 tokens. Thus, in the event that any group member invests, one group member earns 1.25 tokens while the other group members receive 2 tokens. If no group member invests during the 120 second period, each subject earns their endowment, 1 token, for that round. If two participants choose to invest simultaneously, one of them will be randomly selected to invest and pay the cost[11]. This incentive scheme captures the characteristics of a low promotability task as it's a task that makes everyone better off when completed but each individual prefers not to be the individual to invest. At the conclusion of each round, subjects are informed of their earnings and if any, but not which, member of their group invested during the round. Following the experimental conditions, subjects complete an incentivized risk measure and an ex-post survey and are paid their earnings from the experiment in private.

4.3.2 Treatment 2: Asymmetric Volunteer's Dilemma

Treatment 2 is identical to treatment 1 with one important exception, the introduction of heterogeneous costs of investment. At the start of the experiment, participants are told that each participant is randomly assigned a cost of investing that is either high (0.85) or low (0.65) and that each participant's assigned cost will remain the same across all ten rounds. Costs are assigned such

[10] This language is chosen to be neutral and avoid priming subjects (Babcock et al., 2017)
[11] Although subjects are informed of the procedure in the case of a tie, barring server delays, the computer program measures time to the nearest 10ms making the odds of ties unlikely.

that each participant has an equal chance of being assigned either cost but with no restrictions on the distribution of costs in any session. The method in which costs are assigned is explained to participants, but participants are only aware of their own costs. Thus, similar to the incomplete information treatment in Healy and Pate (2018), participants do not know if their group is comprised of individuals with the same or different costs. In an individual round, participants are endowed with a single token and have 120 seconds to choose whether or not to invest. As soon as a group member invests, the round ends and each subject receives an additional token. However, the participant who invested must pay their cost (either 0.65 or 0.85). Thus, in the event that the low promotability task is completed, one group member earns 1.15 or 1.35 tokens while the other group members receive 2 tokens. The rest of the experimental conditions for treatment 2 are identical to treatment 1.

4.3.3 Treatment 3: Asymmetric Volunteer's Dilemma with Managers

This treatment is identical to treatment 3 of Babcock et al. (2017) except that we include heterogeneous costs of investment. After providing informed consent, participants in this session have a photo taken of them that will be used during the experiment. Participants are assigned a cost like in treatment 2 and told that they will be randomly placed into groups of 4 in each of ten rounds. However, the conditions of re-matching are slightly different in this treatment. In the first 5 rounds, participants will be matched into groups such that they will not be re-matched with a group member from any previous round. In the remaining five rounds, matching follows the procedure from the first two treatments. Participants are told that in each round their group of 4 will contain a "red player" and three "green players." Red players cannot invest in the group account but do choose one of the three green players to request to invest.

Requests are solicited using the strategy method: prior to learning the identity of the red player, each participant chooses a group member to request in the event that they are selected as the red player in their group. Prior to the investment stage, each participant views the photos of their three other group members as well as their costs and selects the one they would like to request to complete the task in the event they are selected as the red player in their group. After each participant makes their request, the red player is chosen at random and each subject is informed of their role as well as the group member who was requested to invest. Although the entire group knows which group member was requested, the request is non-binding and imposes no monetary cost on a subject who chooses to ignore it. Following the request stage, the experiment proceeds identically to treatment 2 except that the red player only sees as a waiting screen as they cannot invest. The same payment structure applies to the red player who earns 2 tokens if any green player invests, and 1 token otherwise.

4.3.4 Incentivized Risk Measure and Ex-post Survey

Across all three treatments, relevant individual preferences and demographics are elicited after the completion of round 10 of the game. Risk preferences are measured using Eckel and Grossman (2008) risk elicitation method. In the elicitation, participants select among 1 of 6 50-50 gambles that each increase in variance with the first 5 increasing in expected value as well. To incentivize participant's responses to this elicitation, one subject in each session is selected to receive payment from their Eckel-Grossman choice. Altruism is measured using the procedures set forth by Bekkers and Wilhelm (2010). Social risk preferences are measured by the Domain-Specific Risk-Taking (DOSPERT) scale as proposed by Weber et al. (2002). Agreeableness is

tested by using "The Big-Five Trait Taxonomy" by John and Srivastava (1999). Lastly, subjects respond Zuckerman's sensation seeking scale (Zuckerman, 2007).

Each of the treatments we ran were conducted in ZTREE (Fischbacher, 2007) in the ERL at Texas A&M University. Subjects were recruited from the ERL's subject pool using ORSEE (Greiner, 2015) and no session lasted longer than an hour and a half. Because of the nature of the research question and Babcock et al.'s (2017) results regarding single sex session special attention was paid to recruiting an equal number of men and women for each session. Over the 14 sessions ran across the three treatments, the proportion of women in each session ranged from 40-60%. Although gender was never explicitly mentioned, subjects could see the other participants in their session prior to the start of the experiment. We ran 5 sessions of treatment 1 and 3 sessions of treatment 2 in the fall of 2016 producing 120 and 72 subjects respectively. In the fall of 2018, we ran 6 sessions of treatment 3 producing 120 subjects. In treatment 1 and 2, average earnings were \$14.30 including a \$5 show-up fee. Average earnings did not differ significantly across the two treatments. In treatment 3, average earnings were \$20.30 including a \$10 show-up fee[12].

4.4 Theory and Hypothesis

4.4.1 Symmetric Volunteer's Dilemma (Treatment 1)

Our predictions about behavior across the three treatments is drawn from the game theoretic work of Diekmann (1984, 1985) and the results of Babcock et al. (2017). We first consider the general game theoretic Nash equilibrium predictions for our baseline treatment

[12] Participants earned "tokens" during the experiment and were told that tokens would be converted to cash at the end of the experiment at the rate of 1 token= \$0.50.

(treatment 1) where subjects are in groups of three and must choose between volunteering and earning 0.25 tokens (the 1 token bonus minus the 0.75 token cost of volunteering) or not volunteering and earning either 0 or 1 tokens based on the behavior of other group members. Given that payoffs for one of the options depend on the actions of other group members, it is unsurprising that an optimal strategy depends on the behavior of other group members. Drawing from the game theoretic solutions provided by Diekmann (1985) and the application by Babcock et al. (2017), we identify three Nash equilibria. The most socially efficient but least equitable solution is an asymmetric pure-strategy Nash equilibrium in which one group member volunteers 100% of the time while the other two group members never volunteer leading to a successful investment in each round. The second equilibrium is a symmetric mixed-strategy Nash equilibrium in which each subject volunteers 13.4% of the time leading to an expected investment 35% of the time. The final equilibrium is an asymmetric mixed-strategy Nash equilibrium in which one group member never volunteers while the other two group members each volunteer 25% of the time. This pattern of behavior would lead to a predicted overall investment rate of 44%[13].

To select among these three predictions, we look at previous experimental results. In the study that is directly comparable to our first treatment, Babcock et al. (2017) find substantial evidence that individuals volunteer more than predicted by the symmetric equilibrium. The authors find that groups succeed in producing the collective good 85.4% of the time—a number much larger than each mixed-strategy equilibrium but much lower than the pure-strategy equilibrium.

[13] Following Babcock et al. (2017), if we account for ties than the rate of volunteering become 23.2% in the symmetric mixed equilibrium and 40% in the asymmetric mixed equilibrium predicting the good will be produced 54% and 64% of the time respectively.

We expect to produce a pattern of results similar to Babcock et al. (2017) forming our first hypothesis:

Hypothesis 1: Groups will secure an investment less than the asymmetric pure-strategy Nash equilibrium predicts, but at a much higher rate than that predicted by the mixed-strategy Nash equilibria.

We now draw predictions about how identity will affect the rate of volunteering. We expect to replicate the results from treatment 1 of Babcock et al. (2017) in which the authors find that women volunteer significantly more often than men. The authors suggest a variety of mechanisms that could drive this prediction including women having greater other-regarding preferences (Andreoni and Vesterlund, 2001) or that women are more interested in appearing agreeable (Braiker, 2002). We do not take a stance on the mechanism driving behavior in this treatment but do predict a pattern of results similar to the original study leading to our second hypothesis.

Hypothesis 2: In line with the results of Babcock et al. (2017), Women will volunteer more than men.

4.4.2 Asymmetric Volunteer's Dilemma (Treatment 2)

To draw hypotheses about how the introduction of heterogeneous costs of volunteering will effect behavior, we return to the game theoretic analysis of Diekmann (1984). Diekmann's theoretical work produces a counter-intuitive prediction that in the presence of asymmetric costs, lower cost individuals will be less likely to volunteer. We find this prediction to intractable given its unrealistic implications such as an individual who can't swim volunteering to save a drowning man. In addition, we find this prediction unattractive as previous experimental evidence (Diekmann, 1984; Healy and Pate, 2018) consistently find a much higher rate of volunteering among low cost individuals. Instead, we find Diekmann's alternative suggestive of a "Prominent

solution" (Schelling, 1960) more attractive. In direct support of this alternative prediction, Przepiorka and Diekmann (2018) find high rates of volunteering among low cost individuals and high overall rates of investment in a VD setting with exactly one low cost group member in each group leading to our third hypothesis.

Hypothesis 3: Low cost individuals will volunteer at a much higher rate than their high cost-counterparts.

Turning to how we expect the introduction of asymmetric costs to effect the previously observed gender gap, we consider the theoretical work of Akerloff and Kranton (2000). Akerloff and Kranton (2010) provide anecdotal evidence of the relevance of gender norms in behavior by highlighting the plight of female lawyers who feel caught between meeting the norms associated with being a women and those associated with being a lawyer. Similarly, the inclusion of identity in the standard utility function can aid in the explanation of the gender gap found in Babcock et al. (2017). If being a woman prescribes increased engagement in low promotability tasks, then the gender gap may be explained by identity-driven utility differences. Women may experience a utility loss from not volunteering, while the opposite is true for men. This is also consistent with Selten's (1960) focal point theorem if the symmetric game leads to an identity-driven focal point where women are expected to volunteer. We predict that the introduction of heterogeneous costs of volunteering may alleviate this gender difference in expectations and beliefs about the need to volunteer through providing an even stronger focal point. In line with this prediction, Healy and Pate (2018) find no gender gap in volunteering in their version of the asymmetric VD leading to our fourth hypothesis.

Hypothesis 4: Introducing heterogeneous costs will alleviate the previously found gender gap and result in equal rates of volunteering among men and women.

4.4.3 Asymmetric Volunteer's Dilemma with Managers (Treatment 3)

Our hypotheses regarding the third treatment are driven largely by the logic used to predict behavior in the second treatment. We theorize that the observed gender gap in requests found in Babcock et al. (2017) is driven by gender serving as a weak focal point. We predict that the asymmetric costs acts as a stronger focal point which will drive manager's behavior leading to our fifth hypothesis.

Hypothesis 5: Managers will request low cost rather than high costs individuals to volunteer.

Given that costs are randomly assigned, it follows logically that there will be less opportunities where managers will be choosing among gender as opposed to cost[14] and as a result that the gender gap in requests received will be much smaller leading to our final hypothesis.

Hypothesis 6: Giving managers information about the heterogeneous costs of their employees will alleviate the previously found gender gap in requests to complete a low promotability tasks.

4.5 Results

4.5.1 Symmetric Volunteer's Dilemma (Treatment 1)

Overall, we are unable to replicate the results of Babcock et al. (2017). We find that on average women volunteer 3.09 times over ten rounds while men volunteer 2.89 times producing a difference that is insignificant (two-sided t-test, p=0.67) and roughly one fifth of the original gap. Figure 4.1 presents the probability of volunteering split by gender over ten rounds. Unlike Babcock et al. (2017), we do not find clear separation in the probability of volunteering across gender and

[14] It is possible that a gender effect may still occur in cases where there are multiple low cost or only high cost group members

find rates of volunteering that are close to the midpoint of volunteering between men and women. Our participants manage to have at least one group member volunteer 89.5% of the time, a rate that is higher than the original study.

Figure 4.1: Probability of Volunteering over Ten Rounds (Treatment 1)

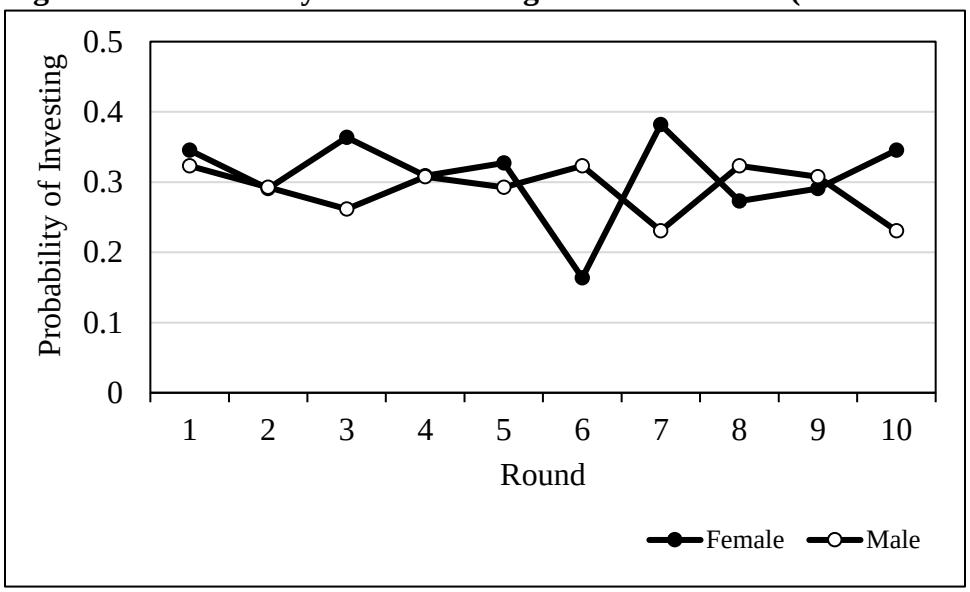


Figure 4.2: Distribution of Total Volunteering over Ten Rounds (Treatment 1)

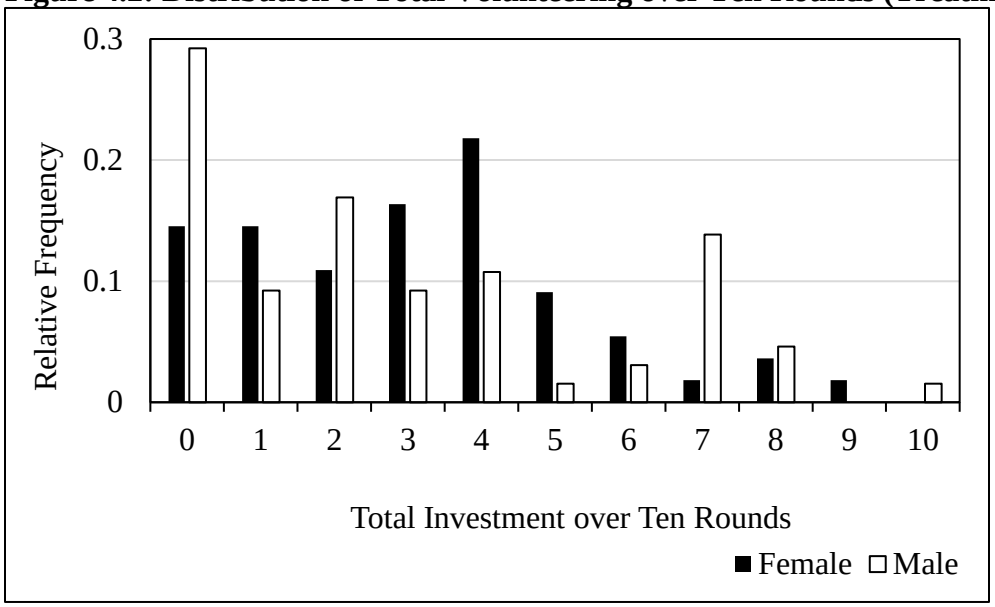


Figure 4.2 presents the distribution of total volunteering over ten rounds. Consistent with Babcock et al. (2017), we find that the distribution of volunteering among males is skewed left while the distribution for women is more normal. The rate of full abstention from volunteering is close to double for men relative to women, a gap that is larger than the one found in the original paper. However, we are unable to reject the hypothesis that the two distributions are drawn from the same population (Kolomogrov-Smirnov test, p=0.48).

Table 4.1: Probit of Volunteering (Treatment 1)

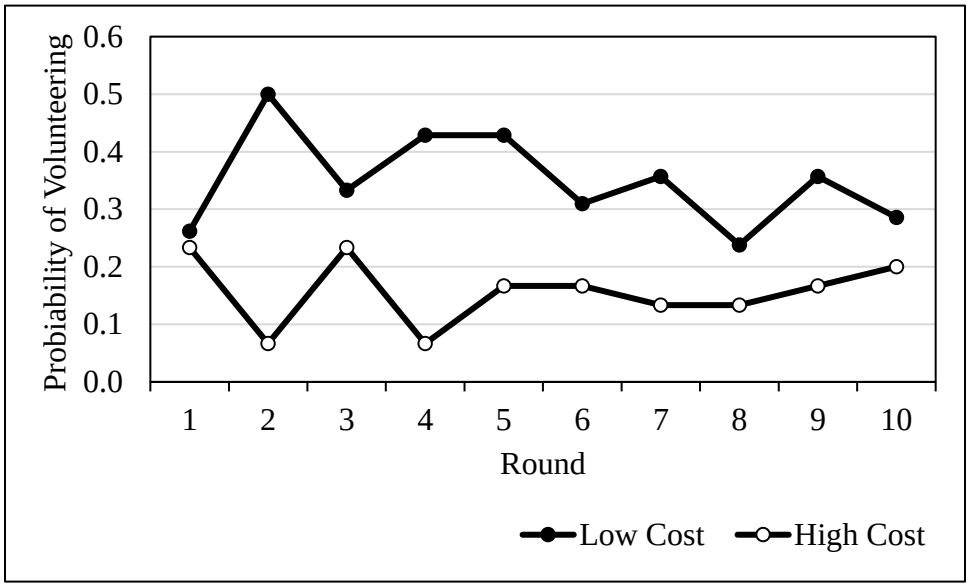
	(1)	(2)
Female	0.020 (0.046)	0.019 (0.045)
Round	-0.003 (0.004)	-0.003 (0.004)
Eckel-Grossman Gamble Choice		-0.001 (0.015)
N	1200	1200

*p<0.10 **p<0.05 ***p<0.01
Dependent variable is individual investment decision (volunteer=1, don't volunteer=0)
Coefficients are marginal effects
Standard errors are clustered at the subject level

Taking a regression approach to our data, table 4.1 displays the results of a probit analysis predicting the probability of volunteering as a function of gender, round, and risk preferences. Standard errors are clustered at the subject level to account for the presence of repeated observation of the same individual. Our model has little predictive power with regards to volunteering behavior and we see a small positive but insignificant effect of being a female on the probability of volunteering. Overall, with the exclusion of the higher rate of non-volunteering among males, we find little to no evidence of a gender difference in volunteering in the symmetric version of the game.

4.5.2 Asymmetric Volunteer's Dilemma (Treatment 2)

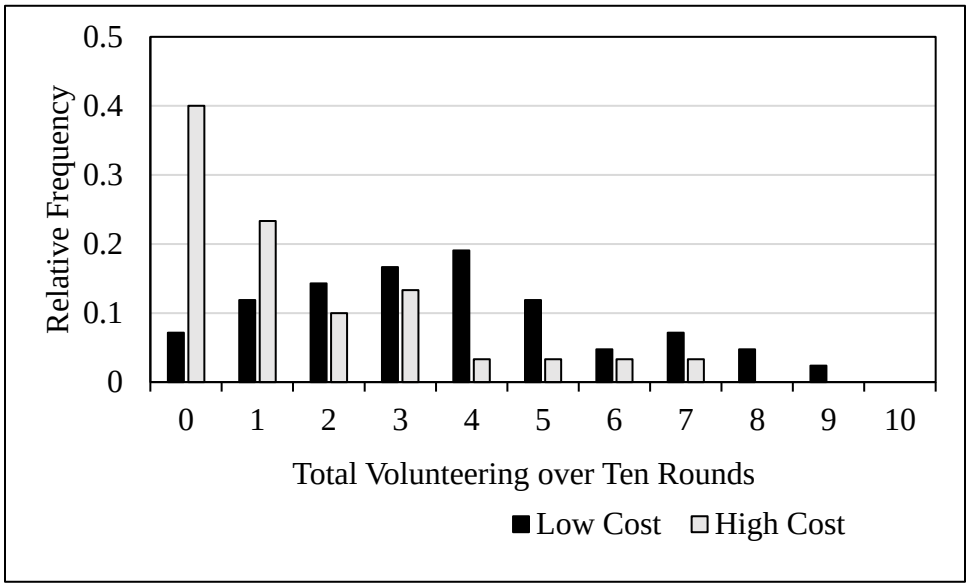
Figure 4.3: Probability of Volunteering by Cost Type (Treatment 2)



Before investigating the gender gap in volunteering in treatment 2, we check whether the introduction of asymmetric costs impacted volunteering behavior. We repeat the same analysis from treatment 1 splitting the data by cost type instead of gender. Looking at the probability of volunteering across rounds (Figure 4.3), there is clear separation across cost types with low cost participants being, on average, over twice as likely to volunteer relative to their high cost counterparts. This separation translates to large gap in overall volunteering with low cost individuals volunteering an average of 3.6 times over ten rounds while their high costs counterparts only volunteer an average of 1.6 times. This difference is significant and almost twice as large in magnitude as the gender gap found in the original study (two-sided t-test, $p<0.001$). Overall, our subjects manage to have at least one group member volunteer 83.3% of the time with a higher rate

of volunteering in the first half of the experiment (92.5% vs. 74.2%). This rate is slightly lower than the original study but the pattern of decay over rounds is similar.

Figure 4.4: Distribution of Total Volunteering by Cost Type (Treatment 2)

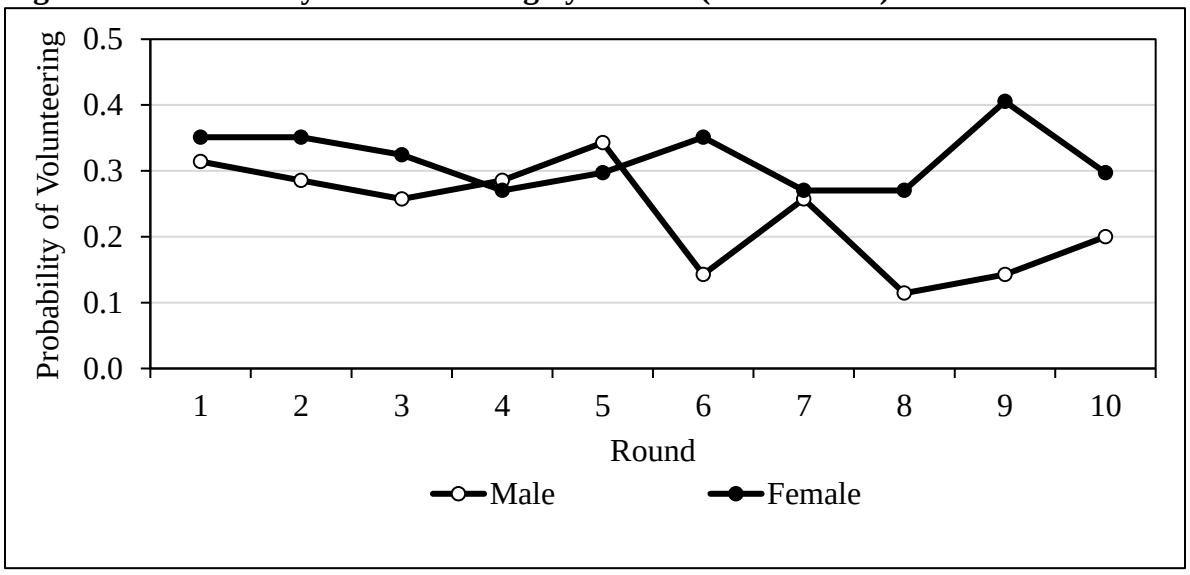


Turning to the distribution of total volunteering over ten rounds (Figure 4.4), we see a pattern that looks similar to the distribution from treatment 1 from Babcock et al. (2017) except that it is based on cost type instead of gender. High cost individuals exhibit a distribution that is skewed left while low cost individuals exhibit a significantly different distribution that is more normal (Kolmogorov-Smirnov test, p=0.002). On this surface, this lends support to our hypothesis that gender norms serve to create a weak focal point in the absence of other signals.

Given our hypotheses and the results from treatment 1, we expect a similar lack of gender gap in treatment 2. Figure 4.5 presents the probability of volunteering across rounds split by gender but pooled across cost type. There is now a larger difference in the probability of volunteering across genders with women volunteering more often than men that becomes larger as the

experiment progresses. This difference translates to a larger overall difference in volunteering than in treatment 1 with men volunteering 2.3 times on average across ten rounds while women volunteer an average of 3.2 times. Although we now find a gap that is similar in magnitude to the original study, we are unable to reject the hypothesis that there is not a significant gender difference because our sample is about half that of the original study (two-sided t-test, p=0.13).

Figure 4.5: Probability of Volunteering by Gender (Treatment 2)



Looking at the distribution of volunteering over 10 rounds (Figure 4.6), we find that the gender disparity in abstention from volunteering is even larger than in treatment 1 or Babcock et al. (2017) with men being over three times as likely to never volunteer across ten rounds. Similar to treatment 1 and the original work, we find again that the distribution for men is skewed left while the distribution for women is more normal. Although the difference in the distribution is more pronounced, we are still unable to reject the claim that the samples are drawn from the same population (Kolmogorov-Smirnov, 0=0.17). Overall, the aggregate results from treatment 2 are

much closer to those from the original study suggesting that our treatment was not only ineffective in reducing the gender gap but may have exacerbated it when considering the results relative to those from treatment 1. However, it is possible that the overall results mask heterogeneity in the interaction of gender and cost type.

Figure 4.6: Distribution of Total Volunteering by Gender (Treatment 2)

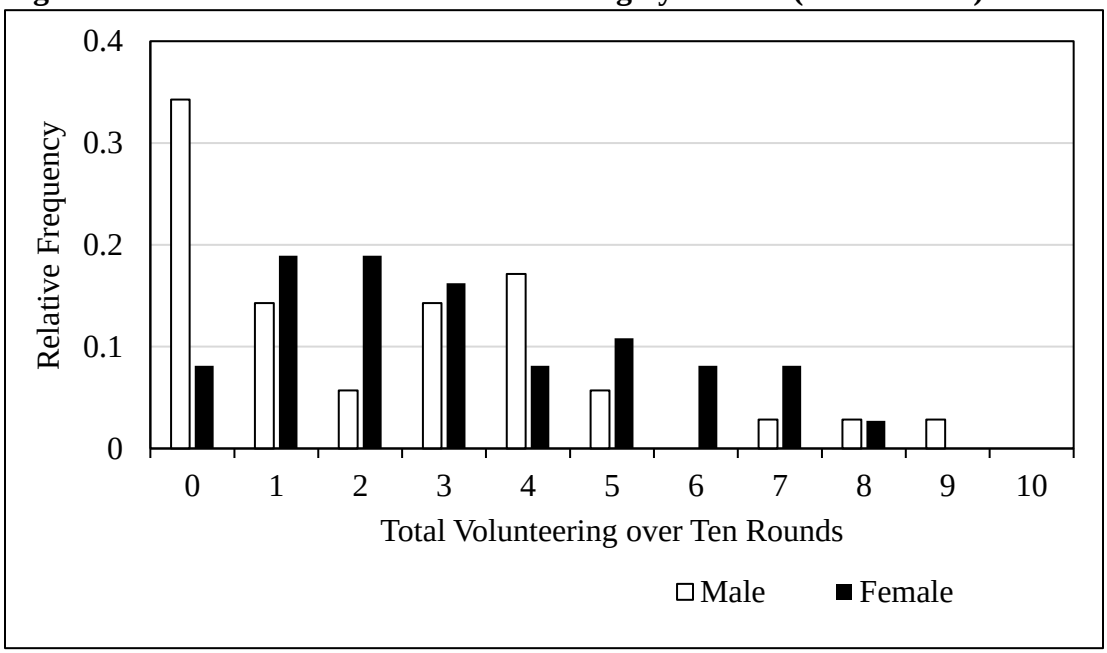


Figure 4.7 presents the probability of investing in each round split by gender among low cost individuals. The rates of volunteering for both men and women are relatively high and overlap often, but the probability of volunteering is still higher among women. We find that, on average, across ten rounds low cost men volunteer 3.3 times while their female counterparts volunteer 4 times (two sided t-test, p=0.34). Relative to the original study, the gap is about two-thirds as large suggesting that our treatment was effective in encouraging greater equality in volunteering among those that should volunteer. However, the gap is larger than the one found in our control treatment

(treatment 1) suggesting that the introduction of asymmetric costs may have exacerbated the gender gap in volunteering.

Figure 4.7: Probability of Volunteering among Low Cost Subjects by Gender (Treatment 2)

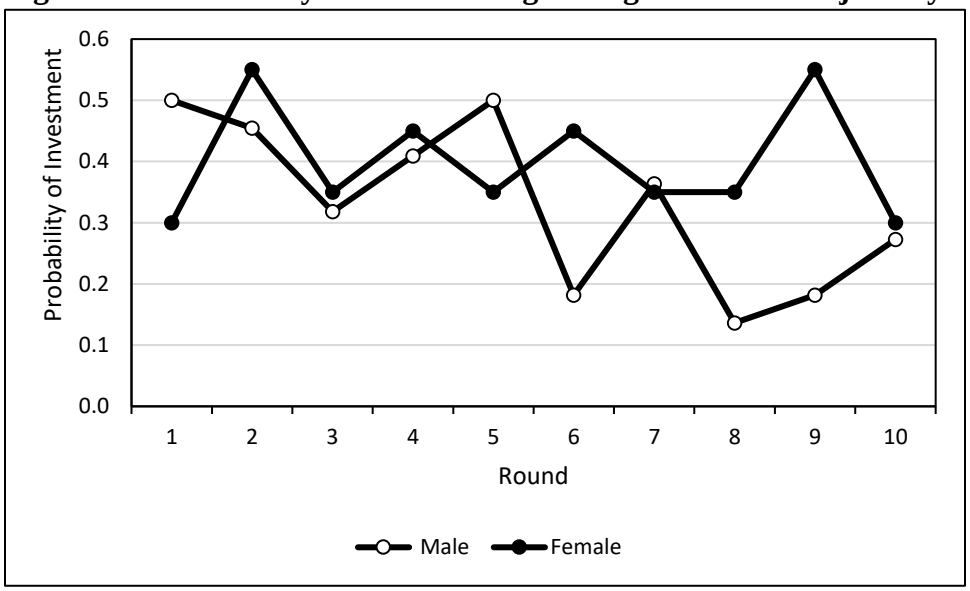
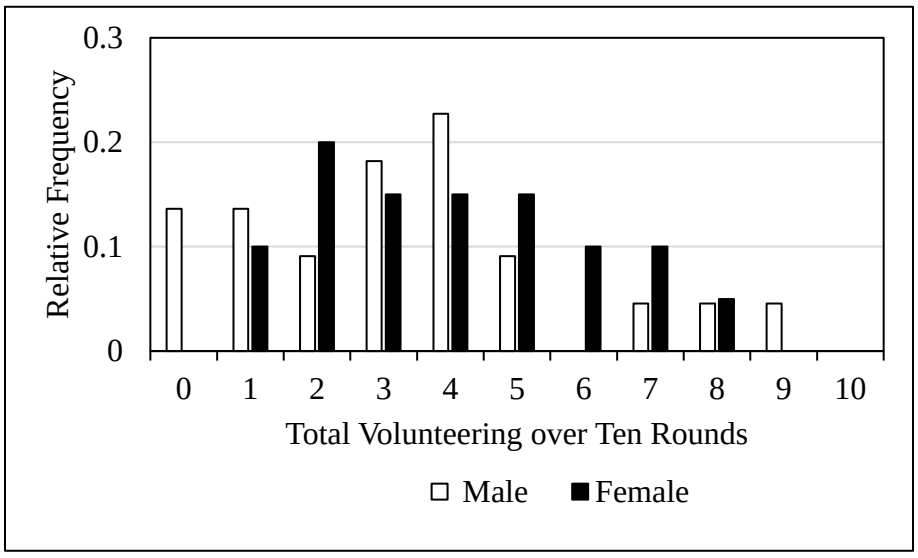
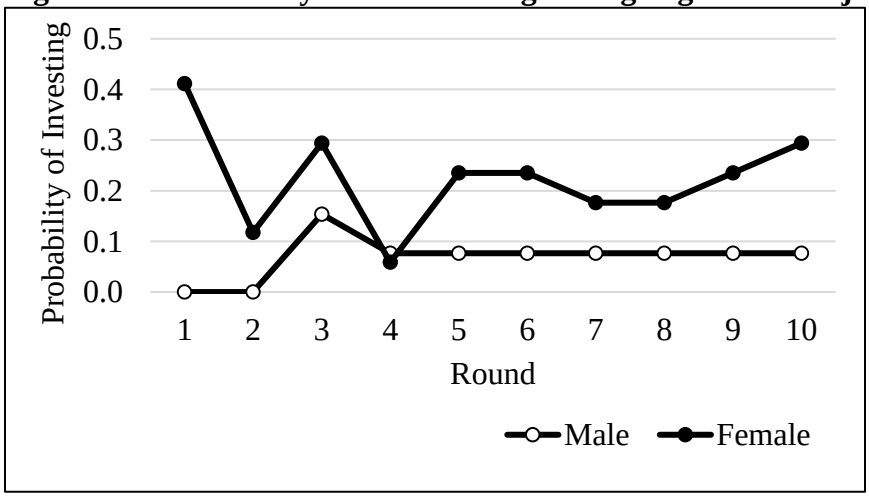


Figure 4.8: Distribution of Total Volunteering among Low Cost Subjects by Gender (Treatment 2)



Looking at the distribution of overall volunteering across ten rounds (figure 4.8), we can again see a pronounced gender difference in abstaining from volunteering with only male low cost participants ever choosing to forgo volunteering in at least one round. However, the overall distribution for both men and women appear relatively normal suggesting that our treatment may have altered the way in which the gender gap in volunteering manifests (Kolmogorov-Smirnov test, p=0.91). Overall, we conclude the relative to the original study, the introduction of heterogeneous costs shrunk the gender gap in volunteering among low cost participants.

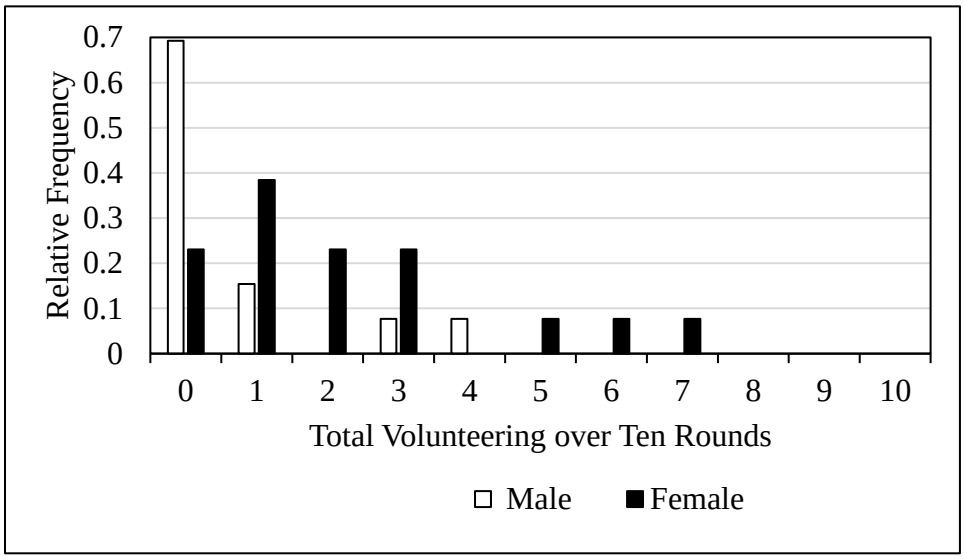
Figure 4.9: Probability of Volunteering among High Cost Subjects by Gender (Treatment 2)



Turning to behavior among high cost participants, we see a stark contrast across genders in the probability of investing (Figure 9). With the exception of round 4, high cost women are always more likely to volunteer and this difference translates into high cost women volunteering at 3 times the rate of high cost men. Over ten rounds, women volunteer an average of 2.24 times while men, on average, volunteer less than once (0.7 times) over ten rounds (two-sided t-test,

p<0.05). The difference within gender among cost types is significant for both men and women (two-sided t-test, p<0.001 and p=0.01 respectively) but the magnitude of the adjustment is much greater for men.

Figure 4.10: Distribution of Total Volunteering among High Cost Subjects by Gender (Treatment 2)



Looking at the distribution of overall volunteering (figure 4.10), we observe that the large gender gap is again driven by the difference in full abstention from volunteering. High cost males are over 3 times as likely to never volunteer and over two-thirds of high cost males choose to never volunteer. Taken together, these results suggest that our treatment exacerbated the gender gap in volunteering among those that should not volunteer suggesting a particularly worrisome pattern in which top-earning females may be more prone to the negative effects of low promotability task completion. Put another way, our results imply that the previously observed gender gap in low

promotability task completion may become even larger as the costs of completion increase creating an even larger divide in labor market outcomes.

To disentangle the effects of cost and gender, we return to our probit analysis with the addition of both gender and type dummies (table 4.2). Consistent with the original paper, we find that across all specifications there is a decay in volunteering over time. In the first specification, we can see that the effect of being a female on the probability of investing is almost identical to the one found in the original study and that the effect of being assigned a low cost is almost twice as strong as the gender effect. When we allow the gender effect to be moderated by cost type (specification 2), we see the much larger gender effect among high cost individuals. Lastly, in the third specification we see that the greater an individual's risk tolerance, the less likely they are to volunteer. Additionally, we see that the gender effect is partly driven by differences in risk tolerance. Overall, we conclude that the introduction of heterogeneous costs did not shrink the previously observed gender gap and may have even exacerbated it. However, this analysis only captures the effect of introducing heterogeneous costs on the supply of low-promotability task completion. In the final section of analysis, we look at how the introduction of heterogeneous costs effects the previously observed gender gap in requests to complete low-promotability tasks.

Table 4.2: Probit of Volunteering (Treatment 2)

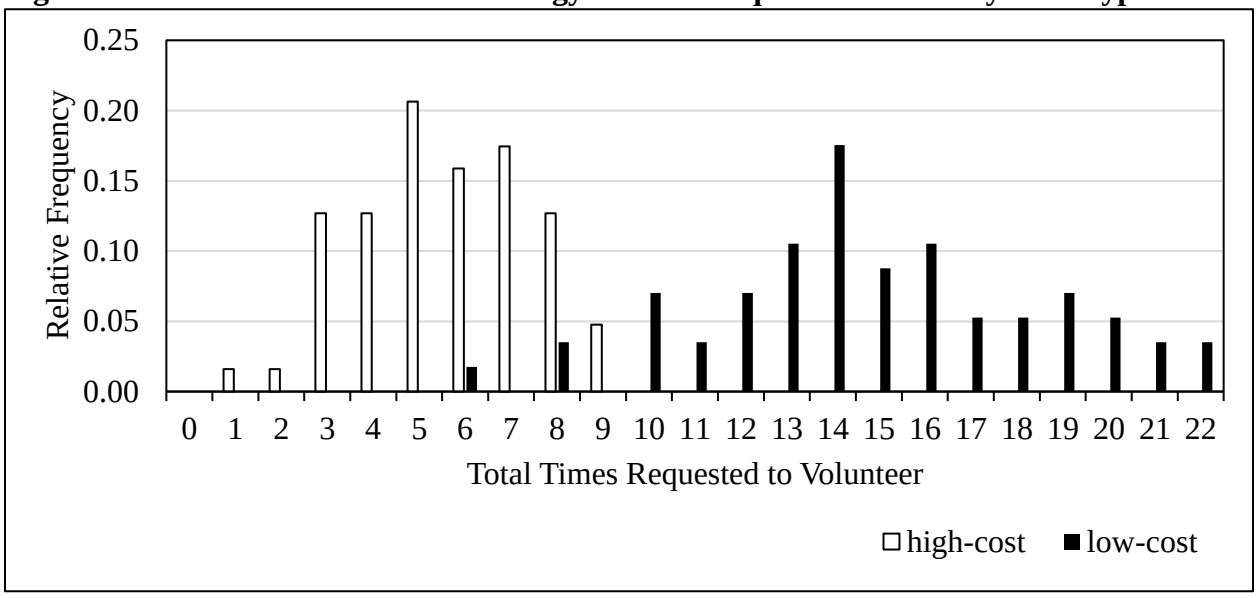
	(1)	(2)	(3)
Female	0.112 (0.050)**	0.228 (0.094)**	0.191 (0.092)**
Low-Cost	0.220 (0.047)***	0.310 (0.079)***	0.295 (0.080)***
Female x Low-Cost		-0.158 (0.096)*	-0.147 (0.093)
Eckel-Grossman Gamble Choice			-0.040 (0.020)**
Round	-0.010 (0.006)*	-0.010 (0.006)*	-0.010 (0.006)*
N	720	720	720

*p<0.10 **p<0.05 ***p<0.01
Dependent variable is individual investment decision (volunteer=1, don't volunteer=0)
Coefficients are marginal effect
Standard errors are clustered at the subject level

4.5.3 Asymmetric Volunteer's Dilemma with Managers (Treatment 3)

In the final treatment, we introduce the concept of managers and ask participants to pick other members of their group to request to volunteer. Surprisingly, unlike the original study, we see a large drop in the rate of volunteering. Across ten rounds, groups manage to have at least one member volunteer 68.5% of the time and the rate does not differ significantly across rounds. In comparison, groups succeed in completing the task 93.5% of the time in the original study. We find that 62.3% of requests are accepted and the rate of acceptance differs only slightly by gender. This rate is lower than the 65.5% acceptance rate from the original study but the major difference across the two samples is among those who volunteer without being requested as only 9% of our participants do so while 14% do so in the original study[15]. Additionally, the rate of accepting requests does not differ significantly across cost types with low cost and high cost individuals accepting 63% and 60% of the time respectively.

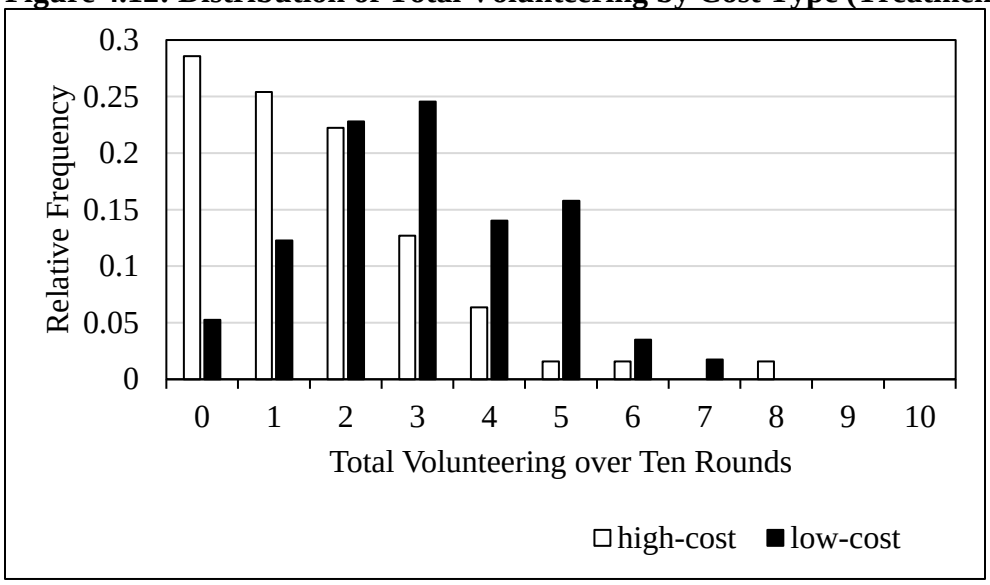
Figure 4.11: Distribution of Total Strategy Method Requests Received by Cost Type



[15] This has a larger effect on overall rate of volunteering as twice as many subjects are not requested

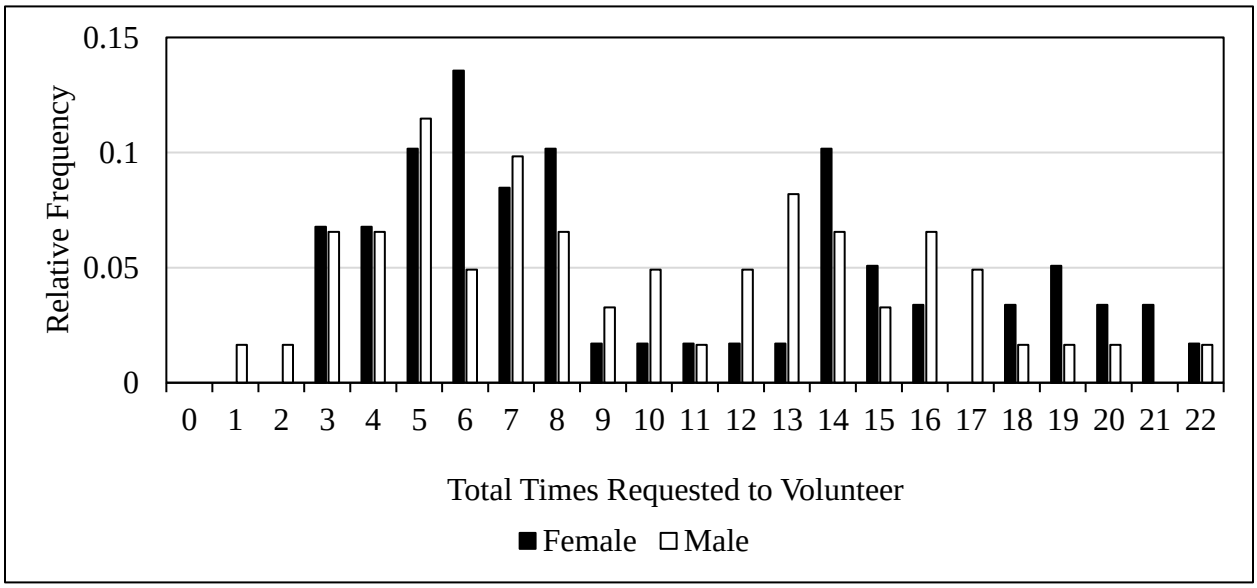
Similar to our analysis to treatment 2, we first investigate whether the introduction of heterogeneous costs had an effect on requests and volunteering behavior. Figure 4.11 presents the distribution of strategy method requests received over ten rounds by cost type. The difference is stark: both samples are relatively normal but the degree of overlap is minimal (Kolmogorov-Smirnov test, $p<0.001$). Managers requested low cost individuals to volunteer an average of 14.9 times while they only request high cost individuals to volunteer an average of 5.6 times (two-sided t-test, $p<0.0001$). Looking to how this affects the rates of volunteering by cost type (figure 4.12), we see that similar to treatment 2 the distribution of volunteering among high cost individuals is skewed left while that of low cost individuals is more normally distributed (Kolmogorov-Smirnov, $p=0.001$). The rate of abstention from volunteering is lower than in treatment 2, and the overall gap in volunteering is smaller but still significant. Low cost individuals volunteer 3 times on average across ten rounds while high cost individuals only volunteer 1.6 times (two-sided t-test, $p<0.0001$). Similar to treatment 2, our pattern of results suggest that the introduction of heterogeneous costs to participants was salient.

Figure 4.12: Distribution of Total Volunteering by Cost Type (Treatment 3)



After confirming that our treatment was effective, we turn to the real question of interest in how the introduction of different costs affected the previously observed gender gap in request received. The distribution of requests received (Figure 4.13) is bi-modal capturing the divisive requesting behavior among cost types. Yet, within gender, we see little difference across the distribution of total requests received (Kolmogorov-Smirnov, p=0.902). On average, we detect only a small difference in requests received as women receive 10.2 requests while men receive 9.8 requests across ten rounds (Two-sided t-test, p=0.69). Thus, we conclude that the introduction of asymmetric costs was effective in culling the previously observed gender gap in requests received to volunteer.

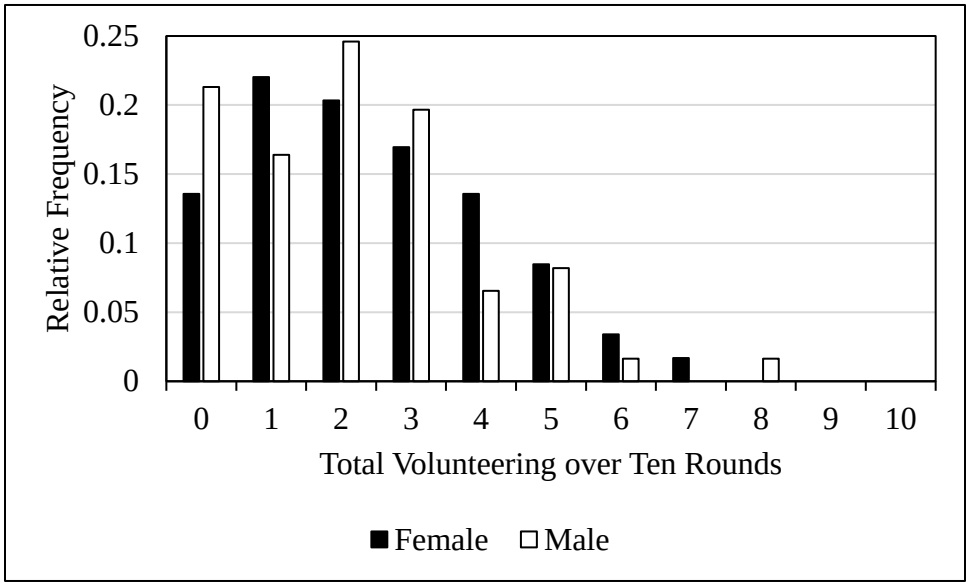
Figure 4.13: Distribution of Total Strategy Method Requests Received by Gender



To see how this translates into actual task completion, figure 4.14 presents the distribution of total volunteering over ten rounds by gender. The distributions for each gender have a slight left-skew, but overall we detect no difference between the distributions (Kolmogorov-Smirnov,

p=0.965). Across 10 rounds women volunteer an average of 2.4 times while men volunteer an average of 2.1 times (two-sided t-test, p=0.39). This difference is driven by a larger but insignificant difference among high cost individuals as high cost women volunteer an average of 3.3 times while their male counterparts only volunteer 2.7 times over ten rounds (two-sided t-test, p=0.19). Overall, it appears that the effect of introducing heterogeneous costs on the gender gap in requests received carried over to task completion.

Figure 4.14: Distribution of Total Volunteering by Gender (Treatment 3)



As a final method of analysis, we repeat the probit analysis from treatment 3. Confirming our earlier analysis, across all 3 specifications we detect no effect of gender. Instead, we find that the only variable with predictive power is whether the participant is low or high cost. Taken together, we conclude that in line with our hypothesis that the introduction of heterogeneous costs creates an effect on requests and volunteering that overpowers the effect of gender.

Table 4.3: Probit of Volunteering (Treatment 3)

	(1)	(2)	(3)
Female	0.032 (0.030)	0.013 (0.050)	0.008 (0.053)
Low-Cost	0.137 (0.030)***	0.119 (0.045)***	0.119 (0.045)***
Female x Low-Cost		0.036 (0.065)	0.035 (0.063)
Eckel-Grossman Gamble Choice			-0.008 (0.012)
Round	0.003 (0.004)	0.003 (0.004)	0.003 (0.004)
N	720	720	720

*p<0.10 **p<0.05 ***p<0.01
Dependent variable is individual investment decision (volunteer=1, don't volunteer=0)
Coefficients are marginal effect
Standard errors are clustered at the subject level

4.6 Conclusion

We find mixed confirmation of our hypotheses. We are unable to replicate the results of Babcock et al. (2017) on which this study is based. However, we find an interesting pattern of results when we introduce asymmetric costs into the previously studied environments. Relative to our baseline treatment, our results imply that introducing explicit cost differences exacerbates the previously found gender gap in low promotability task completion rather than alleviating it as we had hypothesized. We find that the effect of costs differences differs across gender with men being much more responsive to a shift in their cost of completing the task. The resulting effect is that we find a smaller gap relative to the original paper among low-cost participants but a larger gap among high-cost participants. In some sense, this may be worse than finding a null effect as our results imply that the previously observed gap may become even larger as the costs of low promotability task completion increase. This suggest the potential for a much a larger gender disparity in labor market outcomes due to low promotability task completion than previously suggested. However, it may also be viewed as promising that the gap is smaller among those that are meant to volunteer for such tasks. One possible explanation for our pattern of results is that we provided another

weak focal point[16] as opposed to a strong focal point. In support of this possibility, Healy and Pate (2018) detect no gender gap in volunteering when asymmetric costs are introduced, and full information is provided to participants. It is possible that this provided a stronger focal point as there is no uncertainty about if any group contains a low cost member.

We find promising results with regards to requests to complete low promotability tasks. The introduction of explicit opportunity cost differences removes any gender difference in request received, and as a result, in overall completion of low promotability tasks. However, in moving from treatment 2 to treatment 3, we inadvertently changed two variables at once. Beyond the addition of managers, we also unintentionally provided subjects with information about the distribution of costs within their group prior to make volunteering decisions. Consistent with the results of Healy and Pate (2018), it is possible that the change in information rather than the equitable assignment of requests is what produced the lack of gender gap in volunteering. Yet, even so, the result that there is no gender gap in requests in the presence of asymmetric costs is robust. Future work will include an intermediate version of treatment 2 in which information is provided about each of the group member's costs.

Our work provided mixed results on the effect of explicit cost differences on the previously observed gender gap in volunteering to complete low-promotability tasks. On the other hand, we find evidence that the introduction of explicit cost differences alleviates the previously observed gender gap in low promotability task completion. Taken together, our results suggest that it may be the case that gender norms create a weak focal point in the absence of a stronger signal. In the presence of incomplete information about the costs of other group members, there is only a weak individual level focal point on who ought to volunteer. On the other hand, when that

[16] One that may be salient for men but not for women

information is made explicit to managers, a strong focal point is introduced that overpowers the one created by gender norms. However, to be certain of this possibility future work including an intermediate version of treatment 2 in which information is provided about each of the group member's costs is necessary.

5. GENERAL CONCLUSION

My dissertation presents three projects which study the effect of identity on individual preferences and behavior. Each work contains results that illustrate that identity effects how individuals behave in situations that affect themselves as well as situations that affect others. The essays study the effect of general group membership as well as the effect of real social groups and their associated behavioral prescriptions. My work shows that identity is relevant and should be considered in the pursuit of understanding individual heterogeneity in preferences and behavior.

In section 2, we study the ways in which one identity can moderate the effect of another identity on behavior. In particular, we investigate how cultural norms effect the often observed gender gap in competitive preferences. We do so by repeating the procedure of Niederle and Vesterlund (2007) in which subject perform a simple addition task under both a piece rate and competitive payment scheme and then have the opportunity to select which payment scheme they would prefer to have applied to a third performance. Although the original study finds that men are close to twice as likely to pick the competitive payment scheme even though there are no differences in performance, we find a stark contrast within the two populations we studied. In the Hispanic sample which we believe to have cultural norms similar to the original study, we find a gender gap that is similar to the original study. However, in the black sample which we hypothesized has more egalitarian gender norms, we find no gender gap in competitive preferences.

The results of this section suggest that social norms affect individual behavior. Beyond this, our results provide clear evidence that nurture plays a significant role in the expression of competitive preferences as it is unlikely that nature can explain the difference across our

samples. Given other's work showing that competitive preference predict occupational choice (Buser et al., 2012), our work suggests that policymakers should invest in campaigns that aim to shift norms regarding female competitiveness.

My third section turns to studying group behavior more generally by investigating the mechanism underlying in-group bias in trust. Adapting the design introduced by Bohnet and Zeckhauser (2004), we are able to isolate the ways in which a trusting decision differs from a similarly risky decision and how identity effects these considerations. Our results indicate that individuals are not concerned with relative earnings when matched with either members of their in-group or out-group. On the other hand, we do find that subjects are wary to risk a betrayal with either their in-group or out-group by not having trust reciprocated. However, we find this effect is much stronger when paired with members of another group producing the previously observed in-group bias in trust.

Our results add to our understanding of intergroup behavior. Social identification which sorts individuals into different groups leads to discrimination and conflict which are costly from an economic perspective. Our results shed light on the mechanism underlying this behavior and suggest that greater intergroup cooperation may be achieved in the presence of institutions which shield individuals from learning of betrayals.

In my final work, I return to the study the influence of gender identity but do so in a social context. Babcock et al. (2017) provide evidence of a gender gap in low promotability task completion and suggest that this may be a primary cause of the persistence of vertical segregation across genders. We build on this work by testing the importance of gender norms in this context through the introduction of asymmetric costs into the original design. We

hypothesized that gender norms may provide a weak focal point in the absence of a stronger signal of who ought to volunteer to complete low promotability tasks.

Our results provide mixed confirmation of our hypothesis. We fail to replicate the original results but find a gap similar to the original study when asymmetric costs are introduced. We find that the effect of explicit costs differences varies across cost type with low cost individuals exhibiting a smaller gender gap in volunteering than their high cost counterparts. On other hand, we find robust evidence that there is no gender gap in requests received to volunteer when asymmetric costs are introduced. We believe that this is driven by the cost difference providing a stronger focal point that the weak one provided by gender norms. Although it is unclear how making cost differences salient to subjects affects volunteering behavior, our results suggest that making managers aware of other responsibilities of their workers may help to alleviate the empirically observed gender gap in low-promotability task completion.

The work contained within this dissertation provides consistent evidence that individuals are affected by their identity. Individuals recognize and conform to the behavioral prescriptions of the groups to which they belong even when it is costly to do so. Although social norms are difficult to shift, we should all be wary of the subtle but powerful effects they can have on behavior.

REFERENCES

Abbink, Klaus, and Bettina Rockenbach. "Option pricing by students and professional traders: a behavioural investigation." *Managerial and Decision Economics* 27.6 (2006): 497-510.

Aimone, Jason A., and Daniel Houser. "Beneficial betrayal aversion." *PloS one* 6.3 (2011): e17725.

Aimone, Jason A., and Daniel Houser. "What you don't know won't hurt you: a laboratory analysis of betrayal aversion." *Experimental Economics* 15.4 (2012): 571-588.

Aimone, Jason A., and Daniel Houser. "Harnessing the benefits of betrayal aversion." *Journal of Economic Behavior and Organization* 89 (2013): 1-8.

Aimone, Jason A., Daniel Houser, and Bernd Weber. "Neural signatures of betrayal aversion: an fMRI study of trust." *Proceedings of the Royal Society B: Biological Sciences* 281.1782 (2014): 2013-2127.

Akerlof, George A., and Rachel E. Kranton. "Economics and identity." *The Quarterly Journal of Economics* 115.3 (2000): 715-753.

Almås, Ingvild, et al. "Willingness to compete: Family matters." *Management Science* 62.8 (2015): 2149-2162.

Altonji, Joseph G., and Rebecca M. Blank. "Race and gender in the labor market." *Handbook of labor economics* 3 (1999): 3143-3259.

Andersen, Steffen, et al. "Gender, competitiveness, and socialization at a young age: Evidence from a matrilineal and a patriarchal society." *Review of Economics and Statistics* 95.4 (2013): 1438-1443.

Andreoni, James, and Lise Vesterlund. "Which is the fair sex? Gender differences in altruism." *The Quarterly Journal of Economics* 116.1 (2001): 293-312.

Arrow, Kenneth. "The theory of discrimination." *Discrimination in labor markets* 3.10 (1973): 3-33.

Babcock, Linda, Sara Laschever, Michele Gelfand, and Deborah Small. "Nice girls don't ask." *Harvard Business Review* 81.10 (2003): 14-16.

Babcock, Linda, and Sara Laschever. *Women don't ask: Negotiation and the gender divide*. Princeton University Press, 2009.

Babcock, Linda, Maria P. Recalde, Lise Vesterlund, and Laurie Weingart. “Gender differences in accepting and receiving requests for tasks with low promotability.” *American Economic Review* 107.3 (2017): 714-47.

Ball, Sheryl, Catherine C. Eckel, and Maria Heracleous. "Risk aversion and physical prowess: Prediction, choice and bias." *Journal of Risk and Uncertainty* 41.3 (2010): 167-193.

Banuri, Sheheryar, Catherine C. Eckel, and Rick K. Wilson. "Deconstructing nepotism." (2012).

Baron-Cohen, Simon, and Joyce F. Benenson. "Books and arts-essential difference: Men, women and the extreme male brain/essential difference: The truth about the male and female brain." *Nature* 424.6945 (2003): 132-132.

Benjamin, Daniel J., James J. Choi, and A. Joshua Strickland. "Social identity and preferences." *American Economic Review* 100.4 (2010): 1913-28.

Benschop, Yvonne, and Hans Doorewaard. "Covered by equality: The gender subtext of organizations." *Organization Studies* 19.5 (1998): 787-805.

Bergstrom, Ted, Rod Garratt, and Greg Leo. "Let me, or Let George? Motives of competing altruists." (2015).

Bertrand, Marianne, and Kevin F. Hallock. "The gender gap in top corporate jobs." *ILR Review* 55.1 (2001): 3-21.

Bernhard, Helen, Urs Fischbacher, and Ernst Fehr. "Parochial altruism in humans." *Nature* 442.7105 (2006): 912.

Berg, Joyce, John Dickhaut, and Kevin McCabe. "Trust, reciprocity, and social history." *Games and economic behavior* 10.1 (1995): 122-142.

Blau, Francine D., and Lawrence M. Kahn. "The gender wage gap: Extent, trends, and explanations." *Journal of Economic Literature* 55.3 (2017): 789-865.

Blee, Kathleen M., and Ann R. Tickamyer. "Racial differences in men's attitudes about women's gender roles." *Journal of Marriage and the Family* (1995): 21-30.

Bohnet, Iris, and Richard Zeckhauser. "Trust, risk and betrayal." *Journal of Economic Behavior and Organization* 55.4 (2004): 467-484.

Bohnet, Iris, Fiona Greig, Benedikt Herrmann, and Richard Zeckhauser. "Betrayal aversion: Evidence from brazil, china, oman, switzerland, turkey, and the united states." *American Economic Review* 98.1 (2008): 294-310.

Bohnet, Iris, Benedikt Herrmann, and Richard Zeckhauser. "Trust and the reference points for trustworthiness in Gulf and Western countries." *The Quarterly Journal of Economics* 125.2 (2010): 811-828.

Booth, Alison and Patrick Nolen.”Choosing to compete: How different are girls and boys?.” *Journal of Economic Behavior and Organization* 81.2 (2012): 542-555.

Braiker, Harriet. *The disease to please: Curing the people-pleasing syndrome*. McGraw Hill Professional, 2002.

Buser, Thomas. "The impact of the menstrual cycle and hormonal contraceptives on competitiveness." *Journal of Economic Behavior and Organization* 83.1 (2012): 1-10.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. "Gender, competitiveness, and career choices." *The Quarterly Journal of Economics* 129.3 (2014): 1409-1447.

Buser, Thomas, Noemi Peter, and Stefan C. Wolter. "Gender, competitiveness, and study choices in high school: Evidence from Switzerland." *American economic review* 107.5 (2017): 125-30.

Cárdenas, Juan-Camilo, Anna Dreber, Emma von Essen, and Eva Ranehill. "Gender differences in competitiveness and risk taking: Comparing children in Colombia and Sweden." *Journal of Economic Behavior and Organization* 83.1 (2012): 11-23.

Cassar, Alessandra, Feven Wordofa, and Y. Jane Zhang. "Competing for the benefit of offspring eliminates the gender gap in competitiveness." *Proceedings of the National Academy of Sciences* 113.19 (2016): 5201-5205.

Charness, Gary, and Matthew Rabin. "Understanding social preferences with simple tests." *The Quarterly Journal of Economics* 117.3 (2002): 817-869.

Chen, Yan, and Sherry Xin Li. "Group identity and social preferences." *American Economic Review* 99.1 (2009): 431-57.

Chen, Roy, and Yan Chen. "The potential of social identity for equilibrium selection." *American Economic Review* 101.6 (2011): 2562-89.

Daly, Martin, and Margo Wilson, Sex, Evolution, and Behavior, 2nd ed. (Belmont, CA: Wadsworth Publishing Company, 1983).

Dargnies, Marie-Pierre. "Men too sometimes shy away from competition: The case of team competition." *Management Science* 58.11 (2012): 1982-2000.

Dariel, Aurelie, Curtis Kephart, Nikos Nikiforakis, and Christina Zenker. "Emirati women do not shy away from competition: Evidence from a patriarchal society in transition." *Journal of the Economic Science Association* 3.2 (2017): 121-136.

Dave, Chetan, Catherine C. Eckel, Cathleen A. Johnson, and Christian Rojas. "Eliciting risk preferences: When is simple better?." *Journal of Risk and Uncertainty* 41.3 (2010): 219-243.

De Pater, Irene E., Annelies EM Van Vianen, and Myriam N. Bechtoldt. "Gender differences in job challenge: A matter of task allocation." *Gender, Work & Organization* 17.4 (2010): 433-453.

Diekmann, Andreas. "Cooperation in an asymmetric volunteer's dilemma game theory and experimental evidence." *Social Dilemmas and Cooperation*. Springer, Berlin, Heidelberg, 1984. 413-428.

Diekmann, Andreas. "Volunteer's dilemma." *Journal of conflict resolution* 29.4 (1985): 605-610.

Eckel, Catherine C., and Philip J. Grossman. "Managing diversity by creating team identity." *Journal of Economic Behavior and Organization* 58.3 (2005): 371-392.

Eckel, Catherine C., and Philip J. Grossman. "Forecasting risk attitudes: An experimental study using actual and forecast gamble choices." *Journal of Economic Behavior and Organization* 68.1 (2008): 1-17.

Eckel, Catherine C., and Ragan Petrie. "Face value." *American Economic Review* 101.4 (2011): 1497-1513.

Ertac, Seda, and Balazs Szentes. "The effect of information on gender differences in competitiveness: experimental evidence." TÜSİAD-Koç University Economic Research Forum working paper series, 2011.

Falk, Armin, and Christian Zehnder. "A city-wide experiment on trust discrimination." *Journal of Public Economics* 100 (2013): 15-27.

Falk, Armin, and Johannes Hermle. "Relationship of gender differences in preferences to economic development and gender equality." *Science* 362.6412 (2018): eaas9899.

Fershtman, Chaim, and Uri Gneezy. "Discrimination in a segmented society: An experimental approach." *The Quarterly Journal of Economics* 116.1 (2001): 351-377.

Fershtman, Chaim, Uri Gneezy, and Frank Verboven. "Discrimination and nepotism: The efficiency of the anonymity rule." *The Journal of Legal Studies* 34.2 (2005): 371-396.

Fischbacher, Urs. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* 10.2 (2007): 171-178.

Gelfand, Michele J., Jesse R. Harrington, and Joshua Conrad Jackson. "The strength of social norms across human groups." *Perspectives on Psychological Science* 12.5 (2017): 800-809.

Gneezy, Uri, Kenneth L. Leonard, and John A. List. "Gender differences in competition: Evidence from a matrilineal and a patriarchal society." *Econometrica* 77.5 (2009): 1637-1664.

Goette, Lorenz, David Huffman, and Stephan Meier. "The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups." *American Economic Review* 96.2 (2006): 212-216.

Goldin, Claudia, and Cecilia Rouse. "Orchestrating impartiality: The impact of" blind" auditions on female musicians." *American economic review* 90.4 (2000): 715-741.

Gong, Binglin, and Chun-Lei Yang. "Gender differences in risk attitudes: Field experiments on the matrilineal Mosuo and the patriarchal Yi." *Journal of economic behavior and organization* 83.1 (2012): 59-65.

Greiner, Ben. "Subject pool recruitment procedures: organizing experiments with ORSEE." *Journal of the Economic Science Association* 1.1 (2015): 114-125.

Harris, Richard J., and Juanita M. Firestone. "Changes in predictors of gender role ideologies among women: A multivariate analysis." *Sex Roles* 38.3-4 (1998): 239-252.

Healy, Andrew, and Jennifer Pate. "Can teams help to close the gender competition gap?." *The Economic Journal* 121.555 (2011): 1192-1204.

Healy, Andrew J., and Jennifer G. Pate. "Cost asymmetry and incomplete information in a volunteer's dilemma experiment." *Social Choice and Welfare* 51.3 (2018): 465-491.

Hoffman, Moshe, and Uri Gneezy. "Left handed women are more competitive than right handed men: On the biological basis of gender differences in competitiveness." *Work. Pap., Univ. Calif., San Diego* (2010).

Holm, Håkan, and Paul Nystedt. "Intra-generational trust—a semi-experimental study of trust among different generations." *Journal of Economic Behavior and Organization* 58.3 (2005): 403-419.

Hong, Kessely, and Iris Bohnet. "Status and distrust: The relevance of inequality and betrayal aversion." *Journal of Economic Psychology* 28.2 (2007): 197-213.

Hunter, Andrea G., and Sherrill L. Sellers. "Feminist attitudes among African American women and men." *Gender and Society* 12.1 (1998): 81-99.

Idson, Todd L., and Walter Y. Oi. "Workers are more productive in large firms." *American Economic Review* 89.2 (1999): 104-108.

John, Oliver P., and Sanjay Srivastava. "The Big Five trait taxonomy: History, measurement, and theoretical perspectives." *Handbook of personality: Theory and research* (1999): 102-138.

Kane, Emily W. "Racial and ethnic variations in gender-related attitudes." *Annual Review of Sociology* 26.1 (2000): 419-439.

List, John A. "Does market experience eliminate market anomalies?." *The Quarterly Journal of Economics* 118.1 (2003): 41-71.

Li, Sherry Xin, and Tracy Xiao Liu. *Group identity and cooperation in infinitely repeated games*. Mimeo, 2017.

Liu, Elaine M., and Ji Kun Huang. "Risk preferences and pesticide use by cotton farmers in China." *Journal of Development Economics* 103 (2013): 202-215.

Lorber, Judith, and Susan A. Farrell, eds. *The social construction of gender*. Newbury Park, CA: Sage, 1991.

Misra, Joya, Jennifer Hickes Lundquist, and Abby Templer. "Gender, Work Time, and Care Responsibilities Among Faculty 1." *Sociological Forum* 27.2. (2012) Oxford, UK: Blackwell Publishing Ltd.

Mitchell, Sara McLaughlin, and Vicki L. Hesli. "Women don't ask? Women don't say no? Bargaining and service in the political science profession." *PS: Political Science & Politics* 46.2 (2013): 355-369.

Murnighan, J. Keith, Jae Wook Kim, and A. Richard Metzger. "The volunteer dilemma." *Administrative Science Quarterly* (1993): 515-538.

Niederle, Muriel, and Lise Vesterlund. "Do women shy away from competition? Do men compete too much?." *The Quarterly Journal of Economics* 122.3 (2007): 1067-1101.

Niederle, Muriel, Carmit Segal, and Lise Vesterlund. "How costly is diversity? Affirmative action in light of gender differences in competitiveness." *Management Science* 59.1 (2013): 1-16.

Niederle, Muriel. "Gender." In John H. Kagel, Alvin E. Roth Editors (2nd Ed.), *Handbook of Experimental Economics, Volume 2* (2016): 481-563. Princeton, NJ: Princeton University Press.

Noonan, Ryan. "Women in STEM: 2017 Update. ESA Issue Brief# 06-17." *US Department of Commerce* (2017).

Polachek, Solomon William. "Occupational self-selection: A human capital approach to sex differences in occupational structure." *The review of Economics and Statistics* (1981): 60-69.

Przepiorka, Wojtek, and Andreas Diekmann. "Heterogeneous groups overcome the diffusion of responsibility problem in social norm enforcement." *PloS one* 13.11 (2018): e0208129.

Rice University (2018). "College Assignments." Rice.edu. 01 Jun. 2018. Web. 14 Dec 2018. https://dou.rice.edu/residential-colleges/college-assignments

Rose, Amanda J., and Karen D. Rudolph. "A review of sex differences in peer relationship processes: potential trade-offs for the emotional and behavioral development of girls and boys." *Psychological bulletin* 132.1 (2006): 98.

Rosenthal, Abraham Michael. *Thirty-eight witnesses: The Kitty Genovese case*. Open Road Media, 2015.

Schelling, Thomas C. *The strategy of conflict*. Harvard university press, 1980.

Tajfel, Henri, and John C. Turner. "An integrative theory of intergroup conflict." In Mary J. Hatch, Majren Schultz Editors, *Organizational Identity: A Reader* (1979): 56-65. New York, NY: Oxford University Press.

Turner, John C., Rupert J. Brown, and Henri Tajfel. "Social comparison and group interest in ingroup favouritism." *European journal of social psychology* 9.2 (1979): 187-204.

U.S. Census Bureau (2018). "America's Families and Living Arrangements: 2018." Census.gov. 14 Nov. 2018. Web. 15 Dec. 2018. https://www.census.gov/data/tables/2018/demo/families/cps-2018.html

Weber, Elke U., Ann-Renee Blais, and Nancy E. Betz. "A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors." *Journal of behavioral decision making* 15.4 (2002): 263-290.

Weesie, Jeroen. "Asymmetry and timing in the volunteer's dilemma." *Journal of conflict resolution* 37.3 (1993): 569-590.

Wilhelm, Mark Ottoni, and René Bekkers. "Helping behavior, dispositional empathic concern, and the principle of care." *Social Psychology Quarterly* 73.1 (2010): 11-32.

Wilson, Rick K., and Catherine C. Eckel. "Judging a book by its cover: Beauty and expectations in the trust game." *Political Research Quarterly* 59.2 (2006): 189-202.

Wozniak, David, William T. Harbaugh, and Ulrich Mayr. "The menstrual cycle and performance feedback alter gender differences in competitive choices." *Journal of Labor Economics* 32.1 (2014): 161-198.

Zuckerman, Marvin. "The sensation seeking scale V (SSS-V): Still reliable and valid." *Personality and Individual Difference s*43.5 (2007): 1303-1305.

APPENDIX A

Table A-1: Eckel-Grossman Task

Gamble	Payment A	Payment B
1	$10.00	$10.00
2	$14.00	$8.00
3	$18.00	$6.00
4	$22.00	$4.00
5	$26.00	$2.00
6	$30.00	$0.00

Table A-2: Survey Questions

1. What is your gender?

2. How old are you?

3. What is your level of studies?

4. Are you a STEM (Science, Technology, Engineering & Mathematics) major?

5. Are you a member of the ROTC or Corp of Cadets?

6. What is the religion with which you most identify?

7. How important would you say religion is to you?

8. What is the ethnic group with which you most identify?

9. How important would you say your ethnicity is to you?

10. Do you belong to any athletic teams?

11. If so, what is your level of participation?

12. Have you ever attended a single sex institution (grade school, middle school, high school, or college) for at least a year?

13. Did you move to the U.S. before the age of five?

14. How many bathrooms were in your childhood home?

15. Which of the following best describes the neighborhood you grew up in?

16. Were you raised in a household with both parents?

17. Who was the primary decision in your household growing up?

18. What is the highest level of education your father ever completed?

19. How many males (not including yourself) were in your childhood home?

20. How many females (not including yourself) were in your childhood home?

21. Which category best describes your age relative to your siblings?

For the following question, please use a scale from 1 to 6, where 1 means you are "completely unwilling to take risks" and a 6 means you are "very willing to take risks."

22. How do you see yourself; are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

For the following question, please use a scale from 1 to 10, where 1 means you are "lowest relative to most students" and a 10 means you are "highest relative to most students."

23. How intelligent are you relative to the other members of your university?

For each of the following statements, please indicate how much, you as an individual, agree or disagree, with the following statements.

24. I have little control over things that happen to me.

25. I can do just about anything I really set my mind to.

26. I am able to do things as well as most people.

27. Am exacting in my work.

28. I enjoy new and exciting experiences, even if they are a little frightening or unusual.

29. Don't like to draw attention to myself.

30. Don't mind being the center of attention

31. Have you ever been **unfairly** discouraged by a teacher or advisor from continuing your education.

32. People have acted as if they think you are not smart.

33. A working wife feels more useful than one who doesn't hold a job.

34. Girls and boys should be treated the same in school.

35. Competing with boys in school would make a girl unpopular with boys.

36. If there is not enough money for all the children to go to college the boys should get to go instead of the girls.

Table A-3: Repeated Probit Analysis of Task 3 Competition Entry with Beliefs about Relative Performance, Measures of Risk Aversion, and Cultural Controls (Controls Shown)

	(1)	(2)
Female	-0.631 (0.283)**	-0.449 (0.388)
Black	-0.540 (0.228)**	-0.544 (0.241)**
Black x Female	0.847 (0.335)**	0.280 (0.454)
Tournament (Task 2) Performance	0.008 (0.045)	0.016 (0.046)
Tournament (Task2)-Piece Rate (Task1)	0.023 (0.050)	0.021 (0.049)
Guessed Tournament Rank	-0.535 (0.131)***	-0.584 (0.134)***
Eckel-Grossman Gamble Choice	0.138 (0.061)**	0.130 (0.065)**
Self-Reported Risk (NLSY)	0.191 (0.097)**	0.130 (0.065)**
Single Mother		-0.157 (0.387)
Single Mother x Female		-0.186 (0.734)
Single Mother x Female x Black		0.595 (0.737)
Female Head of Household		0.164 (0.316)
Female Head x Female		-0.021 (0.626)
Female Head x Female x Black		0.633 (0.616)
Parental Education Gap (Mother-Father)		0.100(0.069)
Parental Gap x Female		0.092(0.185)
Parental Gap x Female x Black		-0.394(0.199)**
N	244	244

*p<0.10 **p<0.05 ***p<0.01

Standard errors are clustered at the group level

Dependent variable is task 3 choice of compensation scheme (1-tournament 0-piece rate)

Table A-4: Probit of Competition Entry in Task 3 with Black A&M Dummies

	(1)
Female	-1.024 (0.265)***
Black	-0.432 (0.246)*
Black x Female	0.947 (0.331)***
A&M Black	-0.133 (0.411)
A&M Black x Female	-0.068 (0.465)
Tournament (Task 2) Performance	0.098 (0.038)***
Tournament (Task 2)-Piece Rate (Task 1)	0.017 (0.048)
N	244

*p<0.10 **p<0.05 ***p<0.01

Standard errors are clustered at the group level

Dependent variable is task 3 choice of compensation scheme (1-tournament 0-piece rate)

Table A-5. Ordered Probit of Guessed Tournament Rank (Task 2) with Black A&M Dummies

	(1)
Female	0.861 (0.165)***
Black	-0.319 (0.233)
Black x Female	-0.611 (0.289)*
A&M Black	0.731 (0.660)
A&M Black x Female	-0.608 (0.630)
Tournament (Task 2) Performance	-0.301 (0.039)***
Tournament (Task 2)-Piece Rate (Task 1)	-0.028 (0.045)
N	244

*p<0.10 **p<0.05 ***p<0.01

Standard errors are clustered at the group level

Dependent variable is guessed tournament (task 2) rank (1-best, 4-worst)

Table A-6: Repeated Probit Analysis of Task 3 Competition Entry with Beliefs about Relative Performance and Black A&M Dummies

	(1)	(2)
Female	-1.024 (0.265)***	-0.836 (0.273)***
Black	-0.432 (0.246)*	-0.606 (0.236)***
Black x Female	0.947 (0.331)***	0.832 (0.325)***
A&M Black	-0.133 (0.411)	0.137 (0.294)
A&M Black x Female	-0.068 (0.465)	-0.326 (0.415)
Tournament (Task 2) Performance	0.098 (0.038)***	0.002 (0.045)
Tournament (Task 2)-Piece Rate (Task 1)	0.017 (0.048)***	0.010 (0.050)
Guessed Tournament Rank		-0.521 (0.140)***
N	244	244

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is task 3 choice of compensation scheme (1-tournament 0-piece rate)

Table A-7: Probit of Competition Entry in Task 4 with Black A&M Dummies

	(1)
Female	-0.796 (0.246)***
Black	-0.110 (0.235)
Black x Female	0.341 (0.378)
A&M Black	-0.119 (0.400)
A&M Black x Female	0.267 (0.483)
Piece Rate (Task 1) Performance	0.180 (0.039)***
N	244

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is task 4 choice of compensation scheme (1-tournament 0-piece rate)

Table A-8: Ordered Probit of Guessed Piece Rate Rank (Task 1) with Black A&M Dummies

	(1)
Female	0.727 (0.223)***
Black	-0.343 (0.227)
Black x Female	-0.485 (0.288)*
A&M Black	0.380 (0.280)
A&M Black x Female	0.004 (0.313)
Piece Rate Performance	-0.346 (0.035)***
N	244

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is guessed piece rate (task 1) rank (1-best, 4-worst)

Table A-9: Repeated Probit Analysis of Task 4 Competition Entry with Beliefs about Relative Performance and Black A&M Dummies

	(1)	(2)
Female	-0.796 (0.246)***	-0.597 (0.248)**
Black	-0.110 (0.235)	-0.244 (0.260)
Black x Female	0.341 (0.378)	0.201 (0.399)
A&M Black	-0.119 (0.400)	0.014 (0.414)
A&M Black x Female	0.267 (0.483)	0.274 (0.508)
Piece Rate (Task 1) Performance	0.180 (0.039)***	0.062(0.043)
Guessed Piece Rate Rank		-0.549 (0.161)***
N	244	244

*p<0.10 **p<0.05 ***p<0.01
Standard errors are clustered at the group level
Dependent variable is task 4 choice of compensation scheme (1-tournament 0-piece rate).